

**ZWEI MEHRGITTERVERFAHREN
ZUR NUMERISCHEN BERECHNUNG
VON STATIONÄREN STRÖMUNGEN
DURCH PORÖSE MEDIEN MIT FREIEM RAND**

DISSERTATION

zur

Erlangung des Grades eines
Doktors der Naturwissenschaften
der

Abteilung für Mathematik
an der Ruhr-Universität Bochum

von

CHRISTOPH BOLLRATH

aus Lünen

Bochum 1985



Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41546407_0045

**ZWEI MEHRGITTERVERFAHREN
ZUR NUMERISCHEN BERECHNUNG
VON STATIONÄREN STRÖMUNGEN
DURCH PORÖSE MEDIEN MIT FREIEM RAND**

DISSERTATION

zur

Erlangung des Grades eines
Doktors der Naturwissenschaften
der

Abteilung für Mathematik
an der Ruhr-Universität Bochum

von

CHRISTOPH BOLLRATH

aus Lünen

Bochum 1985

Dissertation eingereicht am: 7.2.1985

Tag der mündlichen Prüfung : 23.5.1985

Dekan: Prof. Dr. Dr.h.c. H. Ehlich

Referent: Prof. Dr. D. Braess

Korreferenten: Prof. Dr. H.W. Alt (Bonn)

Prof. Dr. Dr.h.c. H. Ehlich

Inhaltsverzeichnis

Bezeichnungen	I
Einleitung und Übersicht	III
1. Mathematisches Modell und Diskretisierung	1
2. Konvergenz der diskreten Lösung für $h \rightarrow 0$	8
3. Relaxationsverfahren zur Lösung des diskreten Problems	15
4. Der Algorithmus MGDAM	24
5. Der Algorithmus FASDAM mit voller Grob-Gitter-Korrektur	43
6. Numerische Beispiele	59
7. Ergänzende Bemerkungen	76
Anhang	78
Literaturverzeichnis	87
Lebenslauf	90

Bezeichnungen

Reelle Vektoren bezeichnen wir mit unterstrichenen Kleinbuchstaben, Elemente des $\mathbb{R}^2$ auch mit dem großen Buchstaben P.

Von der Feinheit der Diskretisierung unabhängige Konstanten werden stets mit dem Buchstaben C bezeichnet. Dabei ist die Bedeutung eines kleinen c lokaler Natur, sie kann von einem Auftreten zum anderen wechseln.

Alle Integrale sind bzgl. des Lebesgue-Maßes zu verstehen.

Im einzelnen verwenden wir folgende Bezeichnungen:

δ_{ij} ist das Kronecker-Symbol, d.h. $\delta_{ij} := \begin{cases} 1 & \text{falls } i=j, \\ 0 & \text{sonst.} \end{cases}$

I bezeichnet die Einheitsmatrix.

$\overline{p_1 p_2}$ ist der Vektor mit Anfangspunkt p_1 und Endpunkt p_2 .

$\underline{x} \cdot \underline{y}$ steht für das ℓ_2 -Skalarprodukt der Vektoren $\underline{x}$ und $\underline{y}$ aus $\mathbb{R}^n$.

$|\underline{x}|$ bezeichnet die zugehörige Norm, d.h. $|\underline{x}| := (\underline{x} \cdot \underline{x})^{1/2}$.

$\|\underline{x}\|_\infty$ bezeichnet die Maximumnorm von $\underline{x} \in \mathbb{R}^n$.

$\|A\|$ ist die von der euklidischen Norm induzierte Norm (Spektralradiusnorm) der reellen $n \times n$ Matrix A.

$\kappa(A)$ $:= \|A\| \|A^{-1}\|$ ist die Kondition der Matrix A bzgl. der Spektralradiusnorm.

$\overline{M}$ steht für den Abschluß der Menge $M \subset \mathbb{R}^2$.

$\overset{\circ}{M}$ ist das offene Innere der Menge $M \subset \mathbb{R}^2$.

∂M $:= \overline{M} \setminus \overset{\circ}{M}$ ist der Rand von M.

$\chi(M)$ bezeichnet die charakteristische Funktion der Menge M.

card M steht für die Anzahl der Elemente der Menge M.

supp u bezeichnet für eine auf der Menge M definierte Funktion u den abgeschlossenen Träger $\overline{\{x \in M, u(x) \neq 0\}}$.

$\text{span}\{v_1, \dots, v_n\}$

ist das reelle lineare Erzeugnis der Elemente $v_1, \dots, v_n$ eines reellen Vektorraumes.

 $x_n \rightarrow x$

heißt, die Folge $\{x_n\}_{n \in \mathbb{N}}$ reeller Zahlen konvergiert gegen x .

 $u_n \rightarrow u \text{ in } U$

bezeichnet die Norm Konvergenz der Folge $\{u_n\}_{n \in \mathbb{N}}$ in dem normierten Raum U gegen das Element u .

 $u_n \rightarrow u \text{ schwach in } U$

bezeichnet entsprechend die schwache Konvergenz der Folge $\{u_n\}_{n \in \mathbb{N}}$ gegen u , d.h. für alle Elemente v aus dem Dualraum U' von U konvergiert die Folge $\{v(u_n)\}_{n \in \mathbb{N}}$ gegen $v(u)$.

 $v_n \rightarrow v \text{ schwach}^* \text{ in } U'$

bedeutet, daß die Folge $\{v_n\}_{n \in \mathbb{N}}$ in dem Dualraum U' des normierten Raumes U schwach* gegen das Element v konvergiert, d.h. daß für alle $u \in U$ die Beziehung $\lim v_n(u) = v(u)$ gilt.

 $H^m(\Omega)$

bezeichnet für ein Gebiet $\Omega \subset \mathbb{R}^2$ und für $m \in \mathbb{N}$ den üblichen Sobolev-Raum.

 $L^p(\Omega)$

ist für ein Gebiet $\Omega \subset \mathbb{R}^2$ und für $p \in \mathbb{N} \cup \{\infty\}$ der übliche Lebesgue-Raum.

 $|u|_m$

bezeichnet für $u \in H^m(\Omega)$ die $H^m(\Omega)$ -Seminorm, d.h.
 $|u|_m := \left\{ \int_{\Omega} \sum_{|\alpha|=m} |D^\alpha u|^2 \right\}^{1/2}$, wobei für den Multiindex $\alpha \in \mathbb{N} \times \mathbb{N}$ hier $|\alpha|$ als Summe der Komponenten erklärt ist.

 $\|u\|_m$

$:= \left(\sum_{j=0}^m |u|_j^2 \right)^{1/2}$ ist die übliche Norm des Sobolev-Raumes $H^m(\Omega)$.

 $\text{meas } \Omega$

bezeichnet das Lebesgue-Maß von Ω .

 $\int_{\Omega} f$

steht für $\frac{1}{\text{meas } \Omega} \int_{\Omega} f$.

 ∇

bezeichnet den Vektor $(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2})$.

Einleitung und Übersicht

In der vorliegenden Arbeit beschäftigen wir uns mit der numerischen Behandlung freier Randwertprobleme, die den stationären Fluß einer inkompressiblen Flüssigkeit durch ein poröses Medium beschreiben. Das Standardbeispiel ist das Sickern von Wasser durch einen Staudamm.

Die Berechnung dieser Probleme erfolgte zuerst mit heuristischen Verfahren, die iterativ das Lösen eines Randwertproblems auf festem Gebiet mit anschließender Korrektur des Gebietes vorsahen. Einen ausführlichen Überblick über den Einsatz dieser sogenannten "trial free boundary" Methoden gibt C.W. Cryer in [16].

Anfang der siebziger Jahre transformierte C. Baiocchi das Dammproblem mit rechteckigem Damm in eine Variationsungleichung vom Hindernis-Typ (siehe z.B. [6]) und schuf damit die Voraussetzung für die erste befriedigende numerische Behandlung des Problems. Die Berechnung mit Finite Differenzen Methoden wird von C. Baiocchi, V. Comincioli, L. Guerri und G. Volpi in [5] beschrieben. Eine Abschätzung der bei der Diskretisierung mit Finiten Elementen auftretenden Fehler findet man bei R.S. Falk [17]. Ein konvergentes Overrelaxationsverfahren mit Projektion zur Berechnung des diskreten Hindernisproblems wurde von C.W. Cryer schon in [15] vorgestellt.

Die Effektivität der Lösungsverfahren konnte in den letzten Jahren durch Anwendung der Mehrgitter Technik erheblich verbessert werden. Die dem Verfasser in diesem Zusammenhang bekannten Verfahren von A. Brandt, C.W. Cryer [10] und J. Mandel [23] unterscheiden sich im wesentlichen durch die unterschiedliche Übertragung der Nebenbedingung auf das grobe Gitter. W. Hackbusch und H.D. Mittelmann haben das Verfahren aus [10] für nicht-lineare konvexe Minimierungsprobleme mit zweiseitigem Hindernis verallgemeinert [20].

Die Theorie von C. Baiocchi ist jedoch beschränkt auf den Spezialfall eines rechteckigen Gebietes, und auch alle dem Verfasser bekannten Verallgemeinerungen schränken die geometrische Gestalt des Gebietes stark ein.

H.W. Alt [2] und H. Brezis, D. Kinderlehrer, G. Stampacchia [11] entwickelten eine neue (und kompliziertere) Formulierung des Dammproblems, die ohne derartige einschränkende Voraussetzungen auskommt. Dabei werden der hydrostatische Druck u und eine zweite Unbekannte γ als Lösung einer Variationsungleichung beschrieben. In der Arbeit [3] von H.W. Alt findet man die erste numerische Behandlung dieser Variationsungleichung. Die Berechnung der diskreten Lösung erfolgt dort mit Hilfe eines Relaxationsverfahrens vom Jacobi-Typ. Effektivere Rechenverfahren wurden von L.D. Marini und P. Pietra [25] vorgestellt. Neben speziellen Overrelaxationstechniken beschreiben diese Autoren einen Algorithmus, der in jedem Schritt nach einigen Relaxationen das Lösen eines Hilfsproblems in u vom Hindernis-Typ auf dem Teilgebiet mit positivem Druck vorsieht. Die Konvergenz kann nur für eine ineffektive Variante bewiesen werden.

Unabhängig von [25] haben wir einen ähnlichen Algorithmus ("MGDAM") entwickelt, bei dem die auftretenden Hilfsprobleme vom Hindernis-Typ mit Hilfe eines Zweigitterverfahrens gelöst werden. Die Konvergenz unseres Verfahrens läßt sich mit einfachen Argumenten zeigen. Bei der numerischen Behandlung einiger Beispiele zeigte sich der Algorithmus MGDAM den einfachen Relaxationsverfahren (fast immer) überlegen. Jedoch besitzt er nicht die typische Eigenschaft der Mehrgitterverfahren, daß die Konvergenzgeschwindigkeit unabhängig von der Feinheit der Diskretisierung ist.

Daher haben wir einen zweiten Algorithmus ("FASDAM") durch direkte Anwendung der Mehrgitter Technik auf das diskrete Dammproblem entworfen. Dabei können einige Bausteine der bekannten Mehrgitter-Algorithmen zur Behandlung linearer

elliptischer Probleme übernommen werden. Die naive Übertragung führt jedoch zu ineffektiven bzw. divergenten Verfahren. Als kritisch erwies sich vor allem die Restriktion des Defektes. Erst eine Modifikation des durch die Finite Elemente Diskretisierung gegebenen Restriktionsoperators in Randknoten führt zu einem Verfahren, das die für Mehrgitterverfahren typische Effektivität besitzt.

Es folgt eine Übersicht über die Aufteilung der Arbeit:

Im ersten Kapitel wird das kontinuierliche Dammproblem formuliert und anschließend mit Hilfe der Methode der Finiten Elemente diskretisiert.

Das zweite Kapitel beschäftigt sich mit einem Ergebnis von H.W. Alt [3], der Konvergenz der diskreten Lösung gegen eine Lösung des kontinuierlichen Problems.

Die numerische Berechnung des diskreten Problems mit Hilfe von Relaxationsverfahren wird im 3. Kapitel behandelt. Speziell analysieren wir die Eigenschaft, Ober- bzw. Unterlösung zu sein.

In Kapitel 4 definieren wir den Algorithmus MGDAM und zeigen die Konvergenz. Im 5. Kapitel wird der Mehrgitter-Algorithmus FASDAM entworfen. Es handelt sich um ein heuristisches Verfahren; wir zeigen jedoch, daß es wohldefiniert ist in dem Sinne, daß die Fixpunkte des Relaxationsverfahrens aus [3] auch stationäre Punkte von FASDAM sind.

Das Kapitel 6 enthält numerische Beispiele. Wir vergleichen die vorgestellten Verfahren in verschiedenen physikalischen Situationen und bei unterschiedlicher Feinheit der Diskretisierungen. Die Konvergenzeigenschaften werden analysiert. Anschließend berichten wir vom praktischen Einsatz der Programme.

In Kapitel 7 machen wir einige ergänzende Bemerkungen.

Den Anhang dieser Arbeit bilden geplottete Darstellungen von Lösungen zu den Beispielen des Kapitels 6.

Herrn Prof. Dr. D. Braess möchte ich an dieser Stelle für seine ständige Unterstützung meinen besonderen Dank sagen. Die vorliegende Arbeit wurde von ihm angeregt und entstand unter seiner Anleitung. Herrn Prof. Dr. P. Obermann und Herrn Prof. Dr. G. Schmid verdanke ich die Möglichkeit zum praktischen Einsatz der erstellten Programme. Für ihr kollegiales Interesse bedanke ich mich bei FrI. P. Peisker, Herrn J. Karlisch, Herrn Dr. D. Kröner und Herrn Dr. R. Verfürth. Nicht zuletzt gilt mein Dank dem Rechenzentrum der Ruhr-Universität Bochum, das die für diese Arbeit nötigen numerischen Rechnungen ermöglichte.

1. Mathematisches Modell und Diskretisierung

1.1 Formulierung des Problems

Wir betrachten in dieser Arbeit einen vertikalen Schnitt Ω durch ein poröses Medium wie etwa einen Staudamm. Der Rand $\Gamma := \partial\Omega$ läßt sich in drei Teilmengen unterteilen: den Rand zur Luft Γ^0 ,

den Rand zu Wasserreservoirien Γ^+

und den undurchlässigen Teil $\Gamma \setminus (\Gamma^0 \cup \Gamma^+)$.

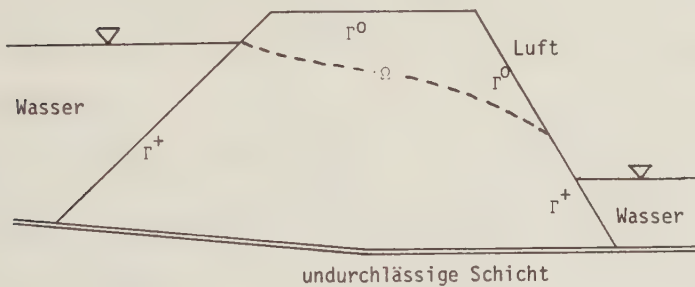


Fig. 1

Die Filterbewegung des Wassers wird durch das Gesetz von Darcy

$$\underline{v} = -a(\nabla u + \underline{e})$$

beschrieben. Dabei bezeichnet $\underline{v}$ die Filtergeschwindigkeit, u den Druck, a die Permeabilitätsmatrix und $\underline{e}$ das äußere Kraftfeld.

Das Wasser wird als inkompressibel betrachtet. Daher besitzt das Gesetz von der Massenerhaltung für eine stationäre Strömung die Gestalt

$$\nabla \cdot \underline{v} = 0.$$

Der Druck u genügt den Randbedingungen

$$u = u^0 \text{ auf } \Gamma^+,$$

$$u \leq u^0 \text{ auf } \Gamma^0.$$

Dabei beschreibt u^0 den Wasserdruck in den Reservoiren bzw. den atmosphärischen Druck.

Die Geschwindigkeit $\underline{v}$ erfüllt die Randbedingungen

$$\underline{v} \cdot \underline{n} = 0 \quad \text{auf} \quad \Gamma \setminus (\Gamma^0 \cup \Gamma^+),$$

$$\underline{v} \cdot \underline{n} \geq 0 \quad \text{auf} \quad \Gamma^0,$$

wobei $\underline{n}$ die nach außen gerichtete Einheitsnormale bezeichnet.

H.W. Alt hat in [2] eine schwache Form des Dammproblems vorgestellt, die im Gegensatz zu den vorher bekannten Formulierungen ohne einschränkende Voraussetzungen an die Daten auskommt. Die in der Realität stark nicht-linear von dem Sättigungsgrad der Strömung abhängige Permeabilität wird dort als Sprungfunktion angesetzt. Zum Nachweis der Existenz einer Lösung wird das so entstandene schwache Problem durch eine Folge von Problemen mit positiver Permeabilität approximiert. Dieser Grenzübergang führt auf folgende Formulierung, bei der a nicht mehr vom Sättigungsgrad abhängt:

$$(1.1) \left\{ \begin{array}{l} \text{Gesucht sind } u \in M(u^0) := \{v \in H^1(\Omega), v = u^0 \text{ auf } \Gamma^+, v \leq u^0 \text{ auf } \Gamma^0\} \\ \text{und } \gamma \in L^\infty(\Omega) \text{ mit } u \geq 0, 0 \leq \gamma \leq 1 \text{ und } \gamma = 1 \text{ auf } \{u > 0\}, \text{ so daß} \\ \int_{\Omega} \nabla(v-u) a(\nabla u + \gamma \underline{e}) \geq 0 \quad \forall v \in M(u^0). \end{array} \right.$$

Der Kapillarbereich konvergiert bei diesem Grenzübergang i.a. nicht gegen eine Hyperfläche. Daher tritt in dem Problem (1.1) neben dem Druck u eine zweite Unbekannte γ auf. Sie beschreibt den Sättigungsgrad der Strömung, d.h. $\{\gamma = 1\}$ kann als gesättigter und $\{\gamma = 0\}$ als trockener Bereich interpretiert werden. Eine in unserem Fall äquivalente Formulierung des Problems findet man auch bei H. Brezis, D. Kinderlehrer und G. Stampacchia [11].

Die Daten sollen in der vorliegenden Arbeit folgende Eigenschaften besitzen: Der Querschnitt Ω ist ein beschränktes, zusammenhängendes Polygon-Gebiet

mit Lipschitz-stetigem Rand Γ . Der Rand zur Luft Γ^0 und der Rand zu Wasserreservoirien Γ^+ sind meßbar und disjunkt, $\Gamma^0 \cup \Gamma^+$ besitzt positives Maß. Die Randwerte des Drucks werden durch eine Funktion $u^0 \in H^1(\Omega) \cap L^\infty(\Omega)$ beschrieben. Speziell sei hier $u^0 = 0$ auf Γ^0 und $u^0 > 0$ auf Γ^+ , d.h. wir vernachlässigen Kapillareffekte und setzen voraus, daß der atmosphärische Druck gleich 0 ist. Der Permeabilitätstensor $a \in L^\infty(\Omega)$ ist strikt elliptisch. Außerdem sollen die Hauptfließrichtungen des Wassers mit den Richtungen der Koordinatenachsen übereinstimmen, a besitzt also Diagonalgestalt. Die Schwerkraft $\underline{e}$ wird durch den konstanten Vektor $(0,1)$ beschrieben.

1.2 Diskretisierung

Wir betrachten eine reguläre Familie konformer Triangulierungen Ω_h von Ω . Dabei sollen alle Dreiecke aus Ω_h eine zu $\underline{e}$ parallele Seite besitzen. Ist diese Forderung nicht erfüllbar, so erhält man nach P. Pietra [26] stabile Diskretisierungen mit Hilfe komplizierterer Finite Elemente Formulierungen, die zusätzlich eine Approximation des Terms $\int_{\Omega} \nabla(v-u) a \underline{e}$ vorsehen.

Bezeichne $P_h := \{P_1, \dots, P_n\}$ die Menge der Knoten von Ω_h . Wir definieren

$$\begin{aligned} P_h^+ &:= P_h \cap \overline{\Gamma^+}, \\ P_h^0 &:= P_h \cap \overline{\Gamma^0} \setminus P_h^+. \end{aligned}$$

Der hydrostatische Druck wird im Raum der stetigen, auf den Dreiecken von Ω_h linearen Finiten Elemente beschrieben. Dabei bezeichne ϕ_h^i das Element der üblichen nodalen Basis mit $\phi_h^i(P_j) = \delta_{ij}$ für alle i, j aus $\{1, \dots, n\}$. Zur Konstruktion des Finite-Elemente-Raumes für γ bestimmen wir zu jedem Punkt $P_i \in P_h$ diejenigen Dreiecke T aus Ω_h ("Up-Wind Dreiecke"), für die

$$\overline{T} \cap \{P_i - \epsilon \underline{e}, 0 < \epsilon < \epsilon_0\} \neq \emptyset \quad \forall \epsilon_0 > 0$$

gilt. Die charakteristische Funktion der Vereinigung dieser Dreiecke bezeichnen wir mit χ_h^i . Die entsprechenden Finite-Elemente-Räume sind

$$S_h := \text{span}\{\phi_h^i, i=1 \dots n\},$$

$$L_h := \text{span}\{\chi_h^i, i=1 \dots n\}.$$

Für $u_h = \sum_{i=1}^n u_h^i \phi_h^i$ schreiben wir statt $(u_h^i)_{i=1}^n$ kurz $\underline{u}_h$ und entsprechend für $\gamma_h = \sum_{i=1}^n \gamma_h^i \chi_h^i$.

Der Variationsungleichung (1.1) entspricht dann das diskrete Problem :

$$(1.2) \left\{ \begin{array}{l} \text{Gesucht sind } \gamma_h \in L_h \text{ und } u_h \in M_h(u_h^0) := \\ \{v_h \in S_h, v_h^i = u_h^{0,i} \text{ für } P_i \in P_h^+, v_h^i \leq u_h^{0,i} \text{ für } P_i \in P_h^0\} \text{ mit} \\ (1.2a) \quad u_h^i \geq 0, \gamma_h^i \in [0,1], u_h^i(1-\gamma_h^i) = 0 \quad \forall P_i \in P_h, \\ \text{so daß} \\ (1.2b) \quad \int_{\Omega} \nabla(v_h - u_h) a_h (\nabla u_h + \gamma_h e) \geq 0 \quad \forall v_h \in M_h(u_h^0). \end{array} \right.$$

Dabei wird die Permeabilität a durch eine punktweise gegen a konvergente Matrixfunktion a_h approximiert. Zur Vereinfachung der Rechnungen fordern wir, daß a_h auf jedem Dreieck aus Ω_h konstant ist und wie a Diagonalgestalt besitzt. Die gleichmäßige Elliptizität darf durch diese Approximation nicht verloren gehen. Außerdem sei $\|a_h\| \leq c$.

Die diskreten Randwerte werden durch ein $u_h^0 \in S_h$ beschrieben mit $u_h^0 \geq 0$, $u_h^0(P) = 0$ falls $P \in P_h^0$ und $u_h^0(P) \geq 0$ falls $P \in P_h^+$. Für $h \rightarrow 0$ konvergiere u_h^0 in $H^1(\Omega)$ gegen u^0 . Setzt man

$$a_{ij}^h := \int_{\Omega} \nabla \phi_h^i a_h \nabla \phi_h^j,$$

$$e_{ij}^h := \int_{\Omega} \nabla \phi_h^i a_h \chi_h^j e,$$

$$A_h := (a_{ij}^h)_{i,j=1}^n,$$

$$E_h := (e_{ij}^h)_{i,j=1}^n,$$

so ist die diskrete Variationsungleichung (1.2b) äquivalent zu:

$$(1.3) \quad (A_h u_h + E_h \gamma_h)^i \begin{cases} = 0 & \text{falls } P_i \in D_h, \\ \leq 0 & \text{falls } P_i \in P_h^0. \end{cases}$$

Dabei bezeichnet D_h die Menge der Knotenpunkte, in denen keine Dirichlet-Randbedingung gestellt ist:

$$D_h := P_h \setminus (P_h^0 \cup P_h^+).$$

Falls keine Verwechslung möglich ist, schreiben wir a_{ij} und e_{ij} statt a_{ij}^h bzw. e_{ij}^h . Die Monotonieeigenschaft der später behandelten Iteration zur Berechnung einer Lösung von (1.2) ergibt sich aus folgendem Lemma, dessen zweiter Teil auch in Kapitel 2 für den Beweis der Konvergenz der diskreten Lösung gegen eine Lösung von (1.1) benötigt wird.

Lemma 1.1

Gelte für alle $x \in \Omega$ die Ungleichung $\kappa(a_h(x)) \leq 3$. Außerdem seien in jedem Dreieck T aus Ω_h alle Winkel kleiner oder gleich $\alpha_T \in (0, \pi)$ mit $\cos \alpha_T = \kappa(a_h|_T - 1) / \kappa(a_h|_T + 1)$. Dann gilt für alle $P_i, P_j \in P_h$:

$$a) \ i \neq j \Rightarrow a_{ij} \leq 0 \quad \text{und} \quad a_{ji} > 0,$$

$$b) \ i \neq j \Rightarrow e_{ij} \leq 0 \quad \text{und} \quad e_{ji} \geq 0.$$

Dieses Lemma entspricht dem Lemma 3.1 in [26]. Der Teil a) verallgemeinert ein altes Ergebnis von P.G. Ciarlet und P.A. Raviart [14].

Beweis:

a) Wir betrachten ein beliebiges Dreieck T aus Ω_h . Seien P_i, P_j, P_k die Eckpunkte des Dreiecks, $\underline{\ell}_i, \underline{\ell}_j$ die Lote von P_i bzw. P_j auf die jeweils gegenüberliegende Seite und β der von ihnen eingeschlossene Winkel:

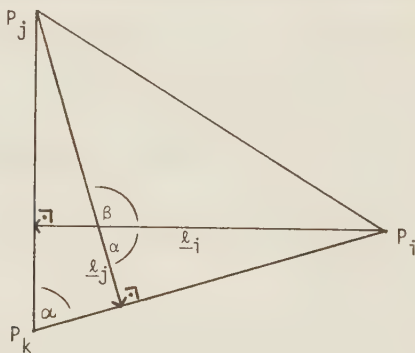


Fig. 2

Dann gilt mit $\alpha := \pi - \beta$ die Beziehung

$$\int_T \nabla \phi_h^i \nabla \phi_h^j = \int_T |\nabla \phi_h^i| |\nabla \phi_h^j| \cos \beta = - \int_T |\nabla \phi_h^i| |\nabla \phi_h^j| \cos \alpha.$$

Bezeichne λ_2 den größten und λ_1 den kleinsten Eigenwert von $a_h|_T$, sei λ das arithmetische Mittel $\frac{1}{2}(\lambda_2 + \lambda_1)$. Der erste Teil der Behauptung a) folgt dann aus der Abschätzung

$$\begin{aligned} \int_T \nabla \phi_h^i a_h \nabla \phi_h^j &= \int_T \nabla \phi_h^i (a_h - \lambda I) \nabla \phi_h^j + \lambda \int_T \nabla \phi_h^i \nabla \phi_h^j \\ &\leq (\|a_h|_T - \lambda I\| - \lambda \cos \alpha) \int_T |\nabla \phi_h^i| |\nabla \phi_h^j| \\ &= \frac{1}{2} (\lambda_2 - \lambda_1 - (\lambda_2 + \lambda_1) \cos \alpha) \int_T |\nabla \phi_h^i| |\nabla \phi_h^j| \leq 0. \end{aligned}$$

Im letzten Schritt wurde von der Voraussetzung

$$(\kappa(a_h|_T) - 1) / (\kappa(a_h|_T) + 1) = (\lambda_2 - \lambda_1) / (\lambda_2 + \lambda_1) \leq \cos \alpha$$

Gebrauch gemacht.

Der zweite Teil der Behauptung a) folgt aus der strikten Elliptizität von a_h .

2. Konvergenz der diskreten Lösung für $h \rightarrow 0$

Wir wollen in diesem Kapitel folgenden Satz zeigen:

Satz 2.1

Sei (u_h, γ_h) Lösung von (1.2). Dann konvergiert mit $h \rightarrow 0$ für eine Teilfolge u_h in $H^1(\Omega)$ gegen ein u und γ_h schwach* in $L^\infty(\Omega)$ gegen ein γ , und (u, γ) ist eine Lösung des Problems (1.1).

Der Beweis dieses Satzes stammt von H.W. Alt (siehe [3], 3.4).

Beweis:

Testet man (1.2b) mit u_h^0 , so erhält man

$$\int_{\Omega} \nabla(u_h^0 - u_h) \cdot a_h \nabla(u_h^0 - u_h) \leq \int_{\Omega} \nabla(u_h^0 - u_h) \cdot a_h (\nabla u_h^0 + \gamma_h e),$$

also

$$|u_h^0 - u_h|_1^2 \leq c |u_h^0 - u_h|_1 \|\nabla u_h^0 + \gamma_h e\|_0 \leq c.$$

Mit den Bezeichnungen

$$\Gamma_h^+ := \{x \in \Gamma, \phi_h^i(x) = 0 \text{ falls } p_i \notin P_h^+\},$$

$$\Gamma_h^0 := \{x \in \Gamma, \phi_h^i(x) = 0 \text{ falls } p_i \notin P_h^0 \cup P_h^+ \setminus \Gamma_h^+\}$$

gilt $\text{meas } \Gamma_h^0 \cup \Gamma_h^+ \geq \text{meas } \Gamma^0 \cup \Gamma^+ - \text{ch.}$ Aufgrund der Poincaré-Ungleichung (siehe etwa Theorem 1.2.1 in [13]) folgt

$$\|u_h^0 - u_h\|_1 \leq c.$$

Mit Hilfe des Einbettungssatzes von Kondrasov folgert man für eine Teilfolge

$$(2.1) \quad \begin{cases} u_h \rightarrow u & \text{in } L^p(\Omega) \text{ für alle } 1 \leq p < \infty; \\ u_h \rightharpoonup u & \text{in } H^1(\Omega) \text{ schwach,} \\ \gamma_h \rightharpoonup \gamma & \text{in } L^\infty(\Omega) \text{ schwach*}. \end{cases}$$

Dann gilt auch $u = u^0$ auf $\Gamma^0 \cup \Gamma^+$, $u \geq 0$ und $0 \leq \gamma \leq 1$ (fast überall).

Sei eine Funktion $v \in M(u^0)$ gegeben. Da sowohl der positive als auch der negative Anteil von $v - u^0$ im $H^1(\Omega)$ -Abschluß der Familie $\{M_h(0), h \in \mathbb{R}^+\}$ liegen, gibt es eine Folge v_h mit

$$(2.2) \quad v_h \rightarrow v \text{ in } H^1(\Omega), \quad v_h \in M_h(u_h^0).$$

Für (u_h, γ_h) , die Lösung des diskreten Problems (1.2), gilt die Beziehung

$$(2.3) \quad \int_{\Omega} \nabla v_h a_h(\nabla u_h + \gamma_h \underline{e}) \geq \int_{\Omega} \nabla u_h a_h(\nabla u_h + \gamma_h \underline{e}).$$

Wir untersuchen die Konvergenz der rechten und der linken Seite von (2.3) getrennt. Für die betrachtete Teilfolge gilt

$$(2.4) \quad \int_{\Omega} \nabla v_h a_h(\nabla u_h + \gamma_h \underline{e}) \rightarrow \int_{\Omega} \nabla v a(\nabla u + \gamma \underline{e});$$

denn aus (2.1), (2.2) und der punktweisen Konvergenz von a_h gegen a schließen wir

$$\begin{aligned} \int_{\Omega} \nabla(v_h - v) a_h(\nabla u_h + \gamma_h \underline{e}) &\rightarrow 0, \\ \int_{\Omega} \nabla v(a_h - a)(\nabla u_h + \gamma_h \underline{e}) &\rightarrow 0, \\ \int_{\Omega} \nabla v a(\nabla(u_h - u) + (\gamma_h - \gamma)\underline{e}) &\rightarrow 0. \end{aligned}$$

Die Untersuchung der rechten Seite von (2.3) erfolgt mit Hilfe der Gleichung

$$\begin{aligned} \int_{\Omega} \nabla u_h a_h \nabla u_h - \int_{\Omega} \nabla u a \nabla u &= \int_{\Omega} \nabla(u_h - u) a_h \nabla(u_h - u) + 2 \int_{\Omega} \nabla(u_h - u) a_h \nabla u + \int_{\Omega} \nabla u (a_h - a) \nabla u \\ &= \int_{\Omega} \nabla(u_h - u) a_h \nabla(u_h - u) + \int_{\Omega} \nabla(2u_h - u)(a_h - a) \nabla u + 2 \int_{\Omega} \nabla(u_h - u) a \nabla u. \end{aligned}$$

Für die betrachtete Teilfolge folgt

$$(2.5) \quad \limsup_{h \rightarrow 0} \int_{\Omega} \nabla u_h a_h \nabla u_h = \int_{\Omega} \nabla u a \nabla u + \limsup_{h \rightarrow 0} \int_{\Omega} \nabla(u_h - u) a_h \nabla(u_h - u).$$

Wegen Lemma 1.1b) gilt

$$\int_{\Omega} \nabla u_h a_h (1 - \gamma_h) \underline{e} = \sum_{1 \leq i, j \leq n} u_h^i e_{ij} (1 - \gamma_h^j) \leq 0$$

und daher (wiederum für die Teilfolge) auch

$$(2.6) \quad \int_{\Omega} \nabla u_h a_h \gamma_h \underline{e} \geq \int_{\Omega} \nabla u_h a_h \underline{e} \rightarrow \int_{\Omega} \nabla u a \underline{e}.$$

Mit (2.3)-(2.6) ist

$$(2.7) \quad \int_{\Omega} \nabla(v-u) a (\nabla u + \gamma \underline{e}) \geq \limsup_{h \rightarrow 0} \int_{\Omega} \nabla(u_h - u) a_h \nabla(u_h - u) + \int_{\Omega} \nabla u a (1 - \gamma) \underline{e}$$

gezeigt.

Damit ist (u, γ) Lösung von (1.1), falls es gelingt für eine Nullmenge N die Beziehung

$$(2.8) \quad \{\gamma < 1\} \setminus N \subseteq \{u=0\}$$

nachzuweisen. Das Einsetzen von $v = u$ in (2.7) liefert dann $u_h \rightarrow u$ in $H^1(\Omega)$.

Zum Nachweis einer solchen Nullmenge sei $\varepsilon > 0$ und $B_\rho(x)$ ein in Ω liegender Kreis mit Radius ρ , so daß die Menge $\{\gamma \leq 1 - \varepsilon\}$ im Mittelpunkt x des Kreises die Dichte 1 besitzt; d.h. es gibt eine Folge $\lambda(r)$ mit $\lambda(r) \rightarrow 0$, so daß

$$\int_{B_r(x)} \chi(\{\gamma \leq 1 - \varepsilon\}) \geq 1 - \lambda(r)$$

ist. Für kleine Radien r gilt dann

$$\int_{B_r(x)} \gamma \leq 1 - \varepsilon + \int_{B_r(x)} \chi(\{\gamma > 1 - \varepsilon\}) \leq 1 - \varepsilon + \lambda(r) \leq 1 - \frac{3}{4} \varepsilon$$

und daher wegen (2.1) für kleine h auch

$$\int_{B_r(x)} \chi(\{\gamma_h < 1\}) = 1 - \int_{B_r(x)} \chi(\{\gamma_h = 1\}) \geq 1 - \int_{B_r(x)} \gamma_h \geq \frac{\varepsilon}{2}.$$

Das nachstehende Lemma 2.2 garantiert die Existenz einer Konstanten C , so daß

$$B_r(x) \cap \{\gamma_h < 1\} \subseteq \{u_h \leq Ch\}$$

ist. Für hinreichend kleine h folgert man sofort

$$\int_{B_r(x)} \chi(\{u_h \leq Ch\}) \geq \int_{B_r(x)} \chi(\{\gamma_h < 1\}) \geq \frac{\varepsilon}{2}.$$

Für $\delta > Ch$ definieren wir

$$\varepsilon_h := \max(\min(2 - \frac{u_h}{\delta}, 1), 0).$$

Dann ist $\varepsilon_h \geq \chi(\{u_h \leq Ch\})$ und es folgt

$$(2.9) \quad \int_{B_r(x)} \varepsilon_h \geq \frac{\varepsilon}{2}.$$

Außerdem gilt

$$(2.10) \quad \int_{B_r(x)} \varepsilon_h \rightarrow \int_{B_r(x)} \max(\min(2 - \frac{u}{\delta}, 1), 0) \leq \int_{B_r(x)} \chi(\{u < 2\delta\}),$$

$$(2.11) \quad \chi(\{u < 2\delta\}) \xrightarrow{\delta \rightarrow 0} \chi(\{u = 0\}) \text{ punktweise monoton fallend.}$$

Mit Hilfe des Satzes von der monotonen Konvergenz schließen wir

$$\int_{B_r(x)} \chi(\{u = 0\}) \geq \varepsilon/2$$

aus (2.9) - (2.11). Die Existenz einer Nullmenge N mit (2.8) wird von Lemma 2.3 garantiert. □

Lemma 2.2

Sei T ein Dreieck aus Ω_h mit $\bar{T} \subset \Omega$. Dann gibt es eine universelle Konstante C , so daß für die Lösung von (1.2) die Ungleichung

$$u_h^j \leq C(u_h^i + h)$$

erfüllt ist. Dabei sind P_i, P_j beliebige Eckpunkte von T .

Das Lemma entspricht dem Lemma 4.3 in [26].

Beweis:

Den dritten Eckpunkt von T bezeichnen wir mit P_k . Für beliebige $P_r, P_s \in P_h$ sei

$$(2.12) \quad a_{rs}^T := \int_T \nabla \phi_h^r a_h \nabla \phi_h^s.$$

Da Ω_h als regulär vorausgesetzt war, folgt

$$\left. \begin{aligned} a_{ss}^T &\leq a_{ss} \leq c \int_{\Omega} |\nabla \phi_h^s|^2 \leq c \\ a_{ss}^T &\geq c \int_T |\nabla \phi_h^s|^2 \geq c \end{aligned} \right\} \quad \forall s \in \{i, j, k\}$$

aus der gleichmäßigen Beschränktheit und Elliptizität von a_h . Also gilt

$$\left. \begin{aligned} a_{ss}^T &\geq ca_{ss} \\ a_{ss}^T / a_{rr}^T &\leq c \end{aligned} \right\} \quad \forall r, s \in \{i, j, k\}.$$

Aus der Definition (2.12) folgern wir

$$a_{ii}^T = -a_{ij}^T - a_{ik}^T, \quad a_{jj}^T = -a_{ij}^T - a_{jk}^T, \quad a_{kk}^T = -a_{ik}^T - a_{jk}^T.$$

Falls $a_{ik}^T \geq a_{ij}^T$ oder $a_{jk}^T \geq a_{ij}^T$ ist, gilt daher $a_{ii}^T \leq -ca_{ij}^T$; andernfalls folgt $a_{ii}^T \leq -2a_{ik}^T$ und $a_{kk}^T \leq -ca_{jk}^T$. Wie im Beweis zu Lemma 1.1a) erkennt man

$0 \geq a_{ij}^T \geq a_{ij}$, also ist im ersten Fall $a_{ii} \leq -ca_{ij}$ gezeigt. Entsprechend gilt im zweiten Fall $a_{ii} \leq -ca_{ik}$ und $a_{kk} \leq -ca_{jk}$. Als Folgerung aus (1.2) und Lemma 1.1a) erhält man

$$-a_{sr} u_h^r = \sum_{\ell \neq r} a_{s\ell} u_h^\ell + \sum_j e_{s\ell} \gamma_h^\ell \leq a_{ss} u_h^s + e_{ss} \quad \forall s, r \in \{i, j, k\}, s \neq r.$$

Damit haben wir

$$u_h^j \leq c(u_h^i + e_{ii}/a_{ii})$$

bzw. im zweiten Fall

$$u_h^k \leq c(u_h^i + e_{ii}/a_{ii}), \quad u_h^j \leq c(u_h^k + e_{kk}/a_{kk})$$

gezeigt. Aus der gleichmäßigen Beschränktheit von a_h und der Regularität der Triangulierung folgt aber

$$e_{rr} = \int_{\Omega} \nabla \phi_h^r a_h \chi_h^r e \leq \left[\int_{\Omega} \nabla \phi_h^r a_h \nabla \phi_h^r \right]^{1/2} \left[\int_{\text{supp} \chi_h^r} e a_h e \right]^{1/2} \leq ch$$

für alle $r \in \{i, j, k\}$. □

Lemma 2.3

Bezeichne B_1^ϵ die Menge der $x \in \Omega$, in denen die Menge $\{\gamma \leq 1-\epsilon\}$ die Dichte 1 besitzt und B_2^ϵ die Menge der $x \in \Omega$, in denen die Dichte von $\{u=0\}$ größer oder gleich $\epsilon/2$ ist. Falls für alle $\epsilon > 0$ die Inklusion $B_1^\epsilon \subseteq B_2^\epsilon$ gilt, dann existiert eine Nullmenge N mit (2.8).

Beweis:

Die Mengen $\{u=0\}$ und $\{\gamma \leq 1-\epsilon\}$ sind Elemente der Borel- σ -Algebra und daher μ -meßbar. Laut Behauptung 2.9.12 in [18], Seite 158, ist für jede μ -meßbare Menge A und für fast alle x die Dichte von A in x gleich $\chi(A)$. Daher gibt

es Nullmengen N_1^ε und N_2^ε , so daß

$$\{\gamma \leq 1-\varepsilon\} = \{\chi(\{\gamma \leq 1-\varepsilon\}) = 1\} \subseteq B_1^\varepsilon \cup N_1^\varepsilon,$$

$$\{u = 0\} = \{\chi(\{u = 0\}) \geq \varepsilon/2\} \supseteq B_2^\varepsilon \setminus N_2^\varepsilon.$$

Mit $N^\varepsilon := N_1^\varepsilon \cup N_2^\varepsilon$ folgt aus $B_1^\varepsilon \subseteq B_2^\varepsilon$ daher

$$\{\gamma \leq 1-\varepsilon\} \subseteq \{u = 0\} \cup N^\varepsilon.$$

Also gilt auch

$$\{\gamma < 1\} = \bigcup_{n \in \mathbb{N}} \{\gamma \leq 1 - \frac{1}{n}\} \subseteq \{u = 0\} \cup N,$$

wobei $N := \bigcup_{n \in \mathbb{N}} N^{\frac{1}{n}}$ ist.

□

3. Relaxationsverfahren zur Lösung des diskreten Problems

3.1 Das Gesamtschrittverfahren nach H.W. Alt

Die Berechnung einer Lösung des diskreten Dammproblems (1.2) erfolgt in [3] mit Hilfe des Relaxationsverfahrens

$$(3.1) \quad (u_h^{(k+1)}, \gamma_h^{(k+1)}) := R_h(u_h^{(k)}, \gamma_h^{(k)}), \quad k = 0, 1, 2, \dots,$$

wobei der Relaxationsoperator mit Projektion

$$R_h : \begin{cases} S_h \times L_h \rightarrow S_h \times L_h \\ (u_h, \gamma_h) \rightarrow R_h(u_h, \gamma_h) := (R_h^u(u_h, \gamma_h), R_h^\gamma(u_h, \gamma_h)) \end{cases}$$

komponentenweise durch

$$(3.2) \quad \begin{cases} (R_h^u(u_h, \gamma_h))^i := \begin{cases} \max\{(C^i(u_h, \gamma_h) - e_{ii})/a_{ii}, 0\} & \text{falls } P_i \in D_h, \\ u_h^{0,i} & \text{falls } P_i \in P_h^0 \cup P_h^+, \end{cases} \\ (R_h^\gamma(u_h, \gamma_h))^i := \begin{cases} \min\{C^i(u_h, \gamma_h)/e_{ii}, 1\} & \text{falls } P_i \in D_h \cup P_h^0 \text{ und } e_{ii} \neq 0, \\ 1 & \text{sonst,} \end{cases} \\ C^i(u_h, \gamma_h) := b_h^i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} u_h^j + e_{ij} \gamma_h^j \end{cases}$$

erklärt ist. Weil später, im Rahmen der Mehrgitter-Behandlung von (1.2), diskrete Variationsungleichungen mit nichttrivialer rechter Seite b_n auftreten, definieren wir den Operator R_h in Abhängigkeit von einem gegebenen Vektor $\underline{b}_h \in \mathbb{R}^n$.

Das diskrete Problem (1.2) besitzt i.a. keine eindeutige Lösung. In einem Punkt $P_i \in D_h$ mit $e_{ii} = 0$, also einem Punkt auf dem unteren, undurchlässigen Rand, wird γ_h etwa nur durch die Verträglichkeitsbedingung (1.2a) beschrieben. Daher ist in diesem Fall die obige Definition $(R_h^\gamma)^i := 1$ günstig. Für jede Iterierte $(u_h^{(k)}, \gamma_h^{(k)})$, $k \geq 1$, mit nichtnegativem $\gamma_h^{(k)}$

ist automatisch in allen Punkten aus P_h die Verträglichkeitsbedingung (1.2a) erfüllt. Der Zusammenhang zwischen (1.2b) und R_h wird durch folgende Lemmata beschrieben:

Lemma 3.1

Für $(u_h, \gamma_h) \in S_h \times L_h$ gelte die Verträglichkeitsbedingung (1.2a).

Es ist genau dann $(u_h, \gamma_h) \geq R_h(u_h, \gamma_h)$ (komponentenweise), wenn

$$(A_h u_h + E_h \gamma_h - b_h)^i \geq 0 \quad \text{für alle } P_i \in D_h,$$

$$(1 - \gamma_h^i)(A_h u_h + E_h \gamma_h - b_h)^i \geq 0 \quad \text{für alle } P_i \in P_h^0,$$

$$u_h^i \geq u_h^{0,i}, \quad \gamma_h^i = 1 \quad \text{für alle } P_i \in P_h^+,$$

sowie die zusätzliche Bedingung

$$e_{ii} = 0 \Rightarrow \gamma_h^i = 1$$

(d.h. falls P_i kein Up-Wind Dreieck besitzt, gilt $\gamma_h^i = 1$) erfüllt ist.

Beweis:

" $\Rightarrow$ " Als erstes betrachten wir einen Punkt $P_i \in D_h$.

Im Fall $\gamma_h^i = 1$ gilt $a_{ii}u_h^i + e_{ii}\gamma_h^i \geq C^i(u_h, \gamma_h)$.

Im Fall $\gamma_h^i < 1$ ist $e_{ii} \neq 0$ und daher $\gamma_h^i \geq \min\{C^i(u_h, \gamma_h)/e_{ii}, 1\}$;

da $u_h^i = 0$ folgt $a_{ii}u_h^i + e_{ii}\gamma_h^i \geq C^i(u_h, \gamma_h)$.

Für einen Punkt $P_i \in P_h^0$ mit $\gamma_h^i = 1$ ist nichts zu zeigen. Der Fall $P_i \in P_h^0$ und $\gamma_h^i < 1$ wird wie oben behandelt.

" $\Leftarrow$ " Sei zunächst wieder $P_i \in D_h$.

Falls $\gamma_h^i = 1$ ist, gilt $a_{ii}u_h^i + e_{ii} \geq C^i(u_h, \gamma_h)$, also

$u_h^i \geq \max\{(C^i(u_h, \gamma_h) - e_{ii})/a_{ii}, 0\}$; denn laut (1.2a) ist u_h^i nicht negativ.

Außerdem impliziert $\gamma_h^i = 1$ sofort $\gamma_h^i \geq (R_h^i(u_h, \gamma_h))^i$.

Im Fall $\gamma_h^i < 1$ gilt $e_{ii}\gamma_h^i \geq C^i(u_h, \gamma_h)$ und $e_{ii} > 0$, also $\gamma_h^i \geq \min\{C^i(u_h, \gamma_h)/e_{ii}, 1\}$. Außerdem folgt $e_{ii} \geq C^i(u_h, \gamma_h)$ und daher $u_h^i \geq \max\{(C^i(u_h, \gamma_h) - e_{ii})/a_{ii}, 0\} = 0$.

Für einen Punkt $P_i \in P_h^0$ wird der Fall $\gamma_h^i < 1$ analog behandelt; im

Fall $\gamma_h^i = 1$ ist nichts zu zeigen. $\square$

Die Projektion in R_h auf den Kegel $\{\gamma_h \leq 1\}$ führt in den P_h^0 -Punkten, in denen ja $u_h = 0$ vorgeschrieben ist, zu einer Art Komplementaritätsbedingung:

Es ist $\gamma_h^i = 1$ oder es gilt $(A_h u_h + E_h \gamma_h - \underline{b}_h)^i \geq 0$.

Eine ähnliche Bedingung tritt in folgendem Lemma bedingt durch die Projektion auf den Kegel $\{u_h \geq 0\}$ in allen Punkten aus $D_h \cup P_h^0$ auf, in denen $e_{ii} = 0$ gilt, und daher $(R_h(u_h, \gamma_h))^i$ nicht von γ_h^i abhängt.

Lemma 3.2

Für $(u_h, \gamma_h) \in S_h \times L_h$ gelte die Verträglichkeitsbedingung (1.2a).

Es ist genau dann $(u_h, \gamma_h) \leq R_h(u_h, \gamma_h)$ (komponentenweise), wenn

$$(A_h u_h + E_h \gamma_h - \underline{b}_h)^i \leq 0 \quad \text{für alle } P_i \in D_h \cup P_h^0 \text{ mit } e_{ii} \neq 0,$$

$$u_h^i (A_h u_h + E_h \gamma_h - \underline{b}_h)^i \leq 0 \quad \text{für alle } P_i \in D_h \cup P_h^0 \text{ mit } e_{ii} = 0,$$

$$u_h^i \leq u_h^{0,i} \quad \text{für alle } P_i \in P_h^0 \cup P_h^+$$

erfüllt ist.

Der Beweis dieses Lemmas ergibt sich wie der Beweis von Lemma 3.1 durch Fallunterscheidung. Da keine neuen Argumente benötigt werden, wollen wir auf ihn verzichten.

Folgerung 3.3

Falls $(u_h, \gamma_h) \in S_h \times L_h$ der Verträglichkeitsbedingung (1.2a) genügt, ist (u_h, γ_h) genau dann Fixpunkt von R_h , wenn

$$\begin{aligned}
& (A_h u_h + E_h \gamma_h - b_h)^i = 0 && \text{für alle } P_i \in D_h \text{ mit } e_{ii} \neq 0, \\
& \left. \begin{aligned} (A_h u_h + E_h \gamma_h - b_h)^i &\geq 0 \\ u_h^i (A_h u_h + E_h \gamma_h - b_h)^i &= 0 \end{aligned} \right\} && \text{für alle } P_i \in D_h \text{ mit } e_{ii} = 0, \\
& \left. \begin{aligned} (A_h u_h + E_h \gamma_h - b_h)^i &\leq 0 \\ (1 - \gamma_h^i) (A_h u_h + E_h \gamma_h - b_h)^i &= 0 \end{aligned} \right\} && \text{für alle } P_i \in P_h^0 \text{ mit } e_{ii} \neq 0, \\
& u_h^i = u_h^{0,i} && \text{für alle } P_i \in P_h^0 \cup P_h^+, \\
& \gamma_h^i = 1 && \text{für alle } P_i \in P_h^+
\end{aligned}$$

und außerdem die zusätzliche Bedingung

$$e_{ii} = 0 \Rightarrow \gamma_h^i = 1$$

erfüllt ist.

Für den Spezialfall $b_h = \underline{0}$ läßt sich die Aussage von Folgerung 3.3 wesentlich vereinfachen; denn infolge von Lemma 1.1 gilt

$$u_h^i = e_{ii} = 0 \Rightarrow (A_h u_h + E_h \gamma_h)^i \leq 0.$$

Folgerung 3.4

In dem Fall $b_h = \underline{0}$ lassen sich folgende Aussagen treffen:

a) Ein Element $(u_h, \gamma_h) \in S_h \times L_h$, das der Verträglichkeitsbedingung (1.2a) genügt, ist genau dann Fixpunkt von R_h , wenn

$$\begin{aligned}
& (A_h u_h + E_h \gamma_h)^i = 0 && \text{für alle } P_i \in D_h, \\
& \left. \begin{aligned} (A_h u_h + E_h \gamma_h)^i &\leq 0 \\ (1 - \gamma_h^i) (A_h u_h + E_h \gamma_h)^i &= 0 \end{aligned} \right\} && \text{für alle } P_i \in P_h^0, \\
& u_h^i = u_h^{0,i} && \text{für alle } P_i \in P_h^0 \cup P_h^+, \\
& \gamma_h^i = 1 && \text{für alle } P_i \in P_h^+
\end{aligned}$$

und außerdem die zusätzliche Bedingung

$$e_{ii} = 0 \Rightarrow \gamma_h^i = 1$$

erfüllt ist.

b) Jeder Fixpunkt von R_h mit $\gamma_h \geq 0$ ist Lösung von (1.2).

Wir bezeichnen ein Element $(u_h^+, \gamma_h^+) \in S_h \times L_h$ als Oberlösung, wenn (u_h^+, γ_h^+) der Verträglichkeitsbedingung (1.2a) genügt, und die Ungleichung $(u_h^+, \gamma_h^+) \geq R_h(u_h^+, \gamma_h^+)$ erfüllt ist. Entsprechend sei der Begriff Unterlösung erklärt. Der Operator R_h ist stetig und monoton:

$$(3.3) \quad (u_h, \gamma_h) \leq (\tilde{u}_h, \tilde{\gamma}_h) \Rightarrow R_h(u_h, \gamma_h) \leq R_h(\tilde{u}_h, \tilde{\gamma}_h).$$

Ausgehend von einer Oberlösung (u_h^+, γ_h^+) und einer Unterlösung (u_h^-, γ_h^-) mit $(u_h^-, \gamma_h^-) \leq (u_h^+, \gamma_h^+)$, liefert die Iteration (3.1) daher zwei monotone Folgen, die gegen (möglicherweise verschiedene) Fixpunkte von R_h konvergieren.

Auf diese Art wird in [3] nicht nur die Existenz einer diskreten Lösung nachgewiesen, sondern auch für etliche Beispiele eine Näherungslösung berechnet. Eine in unserem Fall passende Unterlösung ist offensichtlich

$u_h^- \equiv \gamma_h^- \equiv 0$ (da $b_h = 0$). Durch $\gamma_h^+ \equiv 1$ und $u_h^+ = K - \underline{x}$ mit

$K := \sup_{\underline{x} \in P_h} (u_h^0(\underline{x}) + \underline{x})$ ist eine passende Oberlösung definiert; denn es gilt

$$(A_h u_h^+ + E_h \gamma_h^+)^i = \int_{\Omega} \nabla \phi_h^i a_h (\nabla u_h^+ + \underline{e}) = 0.$$

3.2 Weitere Relaxationsverfahren

In diesem Abschnitt stellen wir einige Ergebnisse aus [25] vor. Der Relaxationsoperator R_h beschreibt ein Gesamtschrittverfahren und kann selbstverständlich im Gauß-Seidel Sinne modifiziert werden. Dazu definieren wir zu jedem Punkt $P_i \in P_h$ einen Operator $R_h^{(i)}$ auf $S_h \times L_h$:

$$(3.4) \quad (R_h^{(i)}(u_h, \gamma_h))^j := \begin{cases} (R_h(u_h, \gamma_h))^i & \text{falls } i=j, \\ (u_h^j, \gamma_h^j) & \text{sonst.} \end{cases}$$

Durch

$$(3.5) \quad (u_h^{(k+1)}, \gamma_h^{(k+1)}) := G_h(u_h^{(k)}, \gamma_h^{(k)}), \quad k = 0, 1, 2, \dots,$$

mit

$$(3.6) \quad G_h(u_h, \gamma_h) := R_h^{(n)} \circ R_h^{(n-1)} \circ \dots \circ R_h^{(1)}(u_h, \gamma_h)$$

wird dann ein Einzelschrittverfahren zur Lösung von (1.2) beschrieben. Wie folgendes Lemma zeigt, erhält G_h die Oberlösungs- bzw. Unterlöseigenschaften.

Lemma 3.5

$$a) \quad (u_h, \gamma_h) \leq R_h(u_h, \gamma_h) \Rightarrow G_h(u_h, \gamma_h) \leq R_h(G_h(u_h, \gamma_h))$$

$$b) \quad (u_h, \gamma_h) \geq R_h(u_h, \gamma_h) \Rightarrow G_h(u_h, \gamma_h) \geq R_h(G_h(u_h, \gamma_h))$$

Beweis:

Wir zeigen a). Aus (3.3) und (3.4) schließen wir

$$(u_h, \gamma_h) \leq R_h(u_h, \gamma_h) \Rightarrow$$

$$(u_h, \gamma_h) \leq R_h^{(i)}(u_h, \gamma_h) \leq R_h(u_h, \gamma_h) \quad \forall i \in \{1, \dots, n\} \Rightarrow$$

$$(u_h, \gamma_h) \leq R_h^{(i)}(u_h, \gamma_h) \leq R_h(u_h, \gamma_h) \leq R_h(R_h^{(i)}(u_h, \gamma_h)) \quad \forall i \in \{1, \dots, n\}.$$

Also gilt

$$(3.7) \quad (u_h, \gamma_h) \leq R_h(u_h, \gamma_h) \Rightarrow R_h^{(i)}(u_h, \gamma_h) \leq R_h(R_h^{(i)}(u_h, \gamma_h)) \quad \forall i \in \{1, \dots, n\}.$$

Wegen (3.6) folgt Behauptung a) durch wiederholtes Anwenden von (3.7). $\square$

Das Einzelschrittverfahren (3.5) konvergiert nicht langsamer als das Gesamtschrittverfahren (3.1) und besitzt dieselben Fixpunkte:

Lemma 3.6

$$a) (u_h, \gamma_h) \leq R_h(u_h, \gamma_h) \Rightarrow R_h(u_h, \gamma_h) \leq G_h(u_h, \gamma_h)$$

$$(u_h, \gamma_h) \geq R_h(u_h, \gamma_h) \Rightarrow R_h(u_h, \gamma_h) \geq G_h(u_h, \gamma_h)$$

$$b) (u_h, \gamma_h) = R_h(u_h, \gamma_h) \Leftrightarrow (u_h, \gamma_h) = G_h(u_h, \gamma_h)$$

Beweis:

a) Sei $(u_h, \gamma_h) \leq R_h(u_h, \gamma_h)$. Wir definieren:

$$(3.8) \quad \begin{cases} (u_h^{(0)}, \gamma_h^{(0)}) := (u_h, \gamma_h), \\ (u_h^{(v+1)}, \gamma_h^{(v+1)}) := R_h^{(v+1)}(u_h^{(v)}, \gamma_h^{(v)}) \quad \text{für } v \in \{0, \dots, n-1\}. \end{cases}$$

Wegen (3.3) und (3.7) gilt dann

$$(u_h^{(v+1)}, \gamma_h^{(v+1)}) \geq (u_h^{(v)}, \gamma_h^{(v)}) \quad \text{für } v \in \{0, \dots, n-1\}.$$

Mit (3.3) folgern wir

$$(G_h(u_h, \gamma_h))^v = (R_h(u_h^{(v-1)}, \gamma_h^{(v-1)}))^v \geq (R_h(u_h, \gamma_h))^v \quad \text{für } v \in \{1, \dots, n\}.$$

Analog zeigt man den zweiten Teil der Behauptung a).

b) Die Implikation $(u_h, \gamma_h) = R_h(u_h, \gamma_h) \Rightarrow (u_h, \gamma_h) = G_h(u_h, \gamma_h)$ folgt aus Teil a).

Sei $(u_h, \gamma_h) = G_h(u_h, \gamma_h)$. Analog zu (3.8) definieren wir eine Folge $(u_h^{(v)}, \gamma_h^{(v)})$ mit Anfangsglied (u_h, γ_h) . Dann ist $(u_h, \gamma_h) = (u_h^{(v)}, \gamma_h^{(v)})$, für alle $v \in \{0, \dots, n\}$ und es gilt

$$(u_h, \gamma_h)^v = (R_h(u_h^{(v-1)}, \gamma_h^{(v-1)}))^v = (R_h(u_h, \gamma_h))^v \quad \text{für alle } v \in \{1, \dots, n\}. \quad \square$$

Bei der Lösung einfacher Probleme benötigt das Verfahren (3.5) nach unseren Erfahrungen weniger als die Hälfte der Rechenzeit des Verfahrens (3.1).

(Siehe auch [25].) Damit erweist sich das Einzelschrittverfahren bei der Behandlung des Dammproblems dem Gesamtschrittverfahren in der gleichen Weise überlegen wie etwa bei der Lösung der diskreten Laplace-Gleichung ([28], Folgerung aus Theorem 4.3).

L.D. Marini und P. Pietra entwerfen in [25] auch kompliziertere Iterationsverfahren, die von der Tatsache Gebrauch machen, daß u_h innerhalb der gesättigten Zone $\{\gamma_h = 1\}$ Lösung von einfacheren linearen Gleichungen bzw. Ungleichungen ist. (Siehe dazu auch Kapitel 4.) Bei der Berechnung des Dammproblems auf einem Rechteck erweist sich das sogenannte "Incorporated SORP" (Incorporated Systematic Overrelaxation with Projection) als das schnellste dieser Verfahren. Dabei wird in allen Punkten $P \in P_h$, in denen $u_h(P) > h$ gilt, mit einem Parameter $\omega \geq 1$ überrelaxiert, in allen anderen Punkten arbeitet dieses Verfahren wie das Einzelschrittverfahren (3.5). Wie bei anderen monotonen Gleichungen zerstört diese Überrelaxation die Ober- bzw. Unterlösungseigenschaft, und es gibt keinen Beweis für die Konvergenz des Verfahrens. Es werden jedoch in [25] zahlreiche Beispiele vorgeführt, bei denen das Incorporated SORP im Gegensatz zu einem Overrelaxationsverfahren mit einheitlichem Parameter stets konvergiert. Zur Beurteilung der Effektivität vergleiche man auch die Ergebnisse in Kapitel 6. Die Beispiele dort wurden mit folgender Variante des Incorporated SORP behandelt:

$$(3.9) \left\{ \begin{array}{l} (\tilde{u}_h^{(k+1)}, i, \tilde{\gamma}_h^{(k+1)}, i) := (R_h(\sum_{j=1}^{i-1} u_h^{(k+1), j} \phi_h^j + \sum_{j=i}^n u_h^{(k), j} \phi_h^j, \\ \sum_{j=1}^{i-1} \gamma_h^{(k+1), j} \chi_h^j + \sum_{j=i}^n \gamma_h^{(k), j} \chi_h^j)) i, \\ \gamma_h^{(k+1), i} := \tilde{\gamma}_h^{(k+1), i}, \\ u_h^{(k+1), i} := \begin{cases} \max\{0, \omega \tilde{u}_h^{(k+1), i} + (1-\omega) u_h^{(k), i}\} & \text{falls } \tilde{u}_h^{(k+1), i} > h, \\ \tilde{u}_h^{(k+1), i} & \text{falls } \tilde{u}_h^{(k+1), i} \leq h. \end{cases} \end{array} \right.$$

In [25] wird statt einer Folge $(u_h^{(k)}, v_h^{(k)})$ von Elementen aus $S_h \times L_h$ eine Folge von Vektoren $\underline{w}_h^{(k)} \in \mathbb{R}^n$ berechnet, die als $\underline{w}_h^{(k)} = \underline{u}_h^{(k)} + \underline{v}_h^{(k)}$ zu interpretieren sind. Dieser zunächst nur formale Unterschied führt zu einer etwas anderen Overrelaxation:

Zu einer aktuellen Näherung $\underline{w}_h^{(k)} = \underline{u}_h^{(k)} + \underline{v}_h^{(k)}$ wird für $i = 1, \dots, n$ zunächst $\tilde{w}_h^{(k+1),i} = \tilde{u}_h^{(k+1),i} + \tilde{v}_h^{(k+1),i}$ bestimmt, wobei $\tilde{u}_h^{(k+1),i}$ und $\tilde{v}_h^{(k+1),i}$ wie in (3.9) definiert sein sollen, und dann gemäß

$$w_h^{(k+1),i} := \begin{cases} \tilde{w}_h^{(k+1),i} & \text{falls } \tilde{w}_h^{(k+1),i} \leq 1+h, \\ \max\{1, \omega \tilde{w}_h^{(k+1),i} + (1-\omega)w_h^{(k),i}\} & \text{falls } \tilde{w}_h^{(k+1),i} > 1+h \end{cases}$$

überrelaxiert.

Bei Vergleichen für ein einfaches Modellproblem mit quadratischem Ω erwies sich die Version (3.9) dem Verfahren mit Overrelaxation der Vektoren $\underline{w}_h = \underline{u}_h + \underline{v}_h$ leicht überlegen.

Bei der Lösung der diskreten Laplace-Gleichung auf einem regelmäßigen Gitter kann man auf bekannte Näherungen des optimalen Overrelaxationsparameters ω_b zurückgreifen. Eine besonders einfach zu berechnende Abschätzung beschreibt folgende Formel, die von R.S. Varga [28] nach einer Idee von P.R. Garabedian entwickelt wurde:

$$\omega_b \leq \frac{2}{1 + 3.015(h^2/\text{meas } \Omega)^{1/2}}.$$

Es wäre zu überlegen, ob ω sowohl für Ober- und Unterlösung getrennt als auch innerhalb des Iterationsprozesses variabel in Abhängigkeit von der Größe der gesättigten Zone festgelegt werden sollte. Jedoch sind uns keine solchen Experimente bekannt.

4. Der Algorithmus MGDAM

Die Verfahren (3.1) und (3.5) sind Relaxationsverfahren mit Projektion. In der Nähe der Lösung wird in nahezu allen Gitterpunkten P mit $\gamma_h(P) = 1$ der Wert von γ_h nicht mehr verändert und u_h gemäß einer üblichen Jacobi- bzw. Gauß-Seidel-Relaxation bestimmt. Diese Verfahren besitzen die Eigenschaft, daß sie vor allem die kurzwelligen Anteile des Fehlers dämpfen. Nach einigen Schritten ist der Fehler glatt, und das Verfahren wird ineffektiv. Von daher liegt der Gedanke nahe, mit Hilfe der Mehrgitter Technik effektivere Algorithmen zu entwickeln. Die Übertragung der Mehrgitter Idee kann keineswegs naiv geschehen. Schon bei der Behandlung der einfachen Variationsungleichungen vom Hindernis-Typ mit mehrstufigen Verfahren treten Schwierigkeiten auf. So lassen sich aufgrund der speziellen Nichtlinearität dieser Probleme die Konvergenzbeweise für lineare elliptische Differentialgleichungen nicht übertragen. (Siehe dazu [23].)

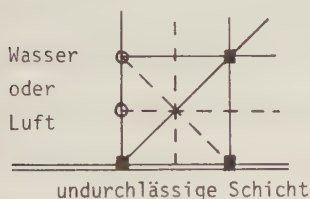
Als einfachster Weg, einen effektiveren Algorithmus zur Lösung von (1.2) zu gewinnen, bietet sich an, die Relaxationsschritte durch zusätzliche Iterationsschritte zu ergänzen, die aus der (näherungsweisen) Berechnung von Hilfsproblemen bestehen, für deren Behandlung Mehrgitterverfahren bekannt sind. Wir bestimmen dabei eine Korrektur nur für den Druck u_h und zwar derart, daß die Einschließung der diskreten Lösung durch Ober- und Unterlösung erhalten bleibt. Bei diesem Vorgehen entsteht wiederum ein konvergentes Verfahren.

Bevor wir den Aufbau des Algorithmus MGDAM näher erläutern, stellen wir die benötigte Gitterstruktur bereit.

4.1 Einführung einer Gitter-Hierarchie

Sei Ω_{h_0} eine Triangulierung von Ω , wie sie in Kapitel 1 beschrieben wurde. Für $k=1, \dots, k_{\max}$ bezeichne Ω_{h_k} eine Verfeinerung von $\Omega_{h_{k-1}}$ mit $h_{k-1}=2h_k$, die ebenfalls alle in Kapitel 1 beschriebenen Eigenschaften besitzen soll. Zusätzlich fordern wir, daß kein Punkt aus $D_{h_{k-1}}$ in Ω_{h_k} einen Nachbarn besitzt, der nicht zu D_{h_k} gehört; d.h. wechselt in $P \in \Omega_{h_{k-1}}$ der Typ der Randbedingung, so soll $P \in P_{h_k}^0 \cup P_{h_k}^+$ gelten:

verboten:



- Gitterlinie des Gitters $\Omega_{h_{k-1}}$
- - - Gitterlinie des Gitters Ω_{h_k}
- Knoten aus $P_{h_k}^0 \cup P_{h_k}^+$
- Knoten aus $D_{h_{k-1}}$

erlaubt:

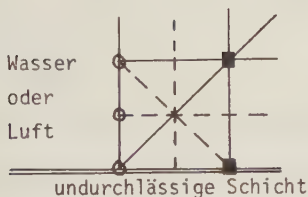


Fig. 4

Im folgenden gehen wir davon aus, daß die Punkte aus P_{h_k} so numeriert sind, daß die ersten $\text{card } P_{h_{k-1}}$ Punkte aus P_{h_k} auch in $P_{h_{k-1}}$ liegen. Es gilt

$$(4.1) \quad S_{h_{k-1}} \subset S_{h_k}, \quad k \in \{1, \dots, k_{\max}\}$$

$$(4.2) \quad L_{h_{k-1}} \subset L_{h_k}, \quad k \in \{1, \dots, k_{\max}\}.$$

Die Relation (4.1) folgt aus der Tatsache, daß Ω_{h_k} Verfeinerung von $\Omega_{h_{k-1}}$ ist.

Zum Beweis von (4.2) sei $h := h_k$, $P_i \in P_{2h}$, $P_j \in P_h$ und $\text{supp } x_h^j \cap \text{supp } x_{2h}^i \neq \emptyset$. Wir haben $\text{supp } x_h^j \subset \text{supp } x_{2h}^i$ zu zeigen. Falls $\text{supp } x_h^j$ nur aus einem Dreieck besteht,

sind wir fertig. Gelte daher für zwei Dreiecke $T_1, T_2 \in \Omega_h$

$$\text{supp } x_h^j = T_1 \cup T_2, \quad T_1 \subseteq \text{supp } x_{2h}^i, \quad T_2 \not\subseteq \text{supp } x_{2h}^i.$$

Wegen $T_2 \cap \text{supp } x_{2h}^i = \emptyset$ folgt $T_1 \cap T_2 \subset \partial(\text{supp } x_{2h}^i)$. Aufgrund der

Konstruktion von x_{2h}^i bedeutet das $T_1 \cap T_2 \subseteq \partial\Omega$ im Widerspruch zur Lipschitzstetigkeit von $\partial\Omega$ ([13], S.12).

Da keine Verwechslung möglich ist, bezeichnen wir sowohl den Operator, der die natürliche Interpolation $S_{2h} \rightarrow S_h$ in den Knotenvektoren beschreibt, als auch den entsprechenden Operator für die Interpolation $L_{2h} \rightarrow L_h$ mit I_{2h}^h .

4.2 Konstruktion eines Hilfsproblems vom Hindernis-Typ

Die erste Komponente R_h^u des Relaxationsoperators (3.2) entspricht formal einer Jacobi-Relaxation für u_h mit Projektion, wie sie üblicherweise bei der Lösung von Hindernisproblemen Verwendung findet. Auf diese Art definiert R_h eine Variationsungleichung in u_h vom Hindernis-Typ. Wir entwerfen in diesem Kapitel einen Algorithmus, der neben Relaxationsschritten gemäß (3.1) bzw. (3.5) auch das Lösen des entsprechenden Hindernisproblems auf dem jeweiligen gesättigten Gebiet $\{\gamma_h = 1\}$ vorsieht. Eine Veränderung von u_h in Punkten der ungesättigten Zone würde möglicherweise die Gültigkeit der Verträglichkeitsbedingung (1.2a) zerstören. Wir wählen daher für $H \in \{h, 2h\}$ folgende Menge als Menge der zulässigen Funktionen:

$$\begin{aligned} \tilde{S}_H(u_h, \gamma_h) &:= \{v_H \in S_H, \quad v_H(P) = 0 \quad \forall P \in P_H^0 \cup P_H^+ \cup (P_H \setminus N_H(\gamma_h)), \\ &\quad v_H(P) + u_h(P) \geq 0 \quad \forall P \in P_H\}, \end{aligned}$$

$$N_H(\gamma_h) := \{P_i \in P_H, \exists v_H \in S_H [v_H(P)(1-\gamma_h(P)) = 0 \forall P \in P_H \text{ und } v_H(P_i) > 0\}.$$

Im Fall $H = h$ entspricht $N_H(\gamma_h)$ der Menge derjenigen Knoten aus P_H , in denen $\gamma_h = 1$ gilt. Außerdem ist $N_H(\gamma_h)$ so gewählt, daß auch im Fall $H = 2h$ mit (u_h, γ_h) alle Elemente aus $(u_h + \tilde{S}_H(u_h, \gamma_h), \gamma_h)$ der Verträglichkeitsbedingung (1.2a) genügen.

Zu gegebenem $(u_h, \gamma_h) \in S_h$ betrachten wir also folgende Variationsungleichung vom Hindernis-Typ:

$$(4.3) \quad \left\{ \begin{array}{l} \text{Gesucht ist } v_h \in \tilde{S}_h(u_h, \gamma_h), \text{ so daß} \\ (\underline{v}_h - v_h)(A_h(\underline{u}_h + v_h) + E_h \gamma_h) \geq 0 \quad \forall \underline{v}_h \in \tilde{S}_h(u_h, \gamma_h). \end{array} \right.$$

Das folgende Lemma besagt, daß die Konvergenz des Gesamtschrittverfahrens (3.1) bzw. des Einzelschrittverfahrens (3.5) nicht zerstört wird, wenn Iterationsschritte der Form $(u_h, \gamma_h) \leftarrow (u_h + v_h, \gamma_h)$, v_h Lösung von (4.3), eingeschoben werden.

Lemma 4.1

Es genüge $(u_h, \gamma_h) \in S_h \times L_h$ der Verträglichkeitsbedingung (1.2a), und es sei v_h Lösung von (4.3). Bezeichne (u_h^*, γ_h^*) einen Fixpunkt des gemäß (3.2) (mit $\underline{b}_h = \underline{0}$) definierten Relaxationsoperators R_h . Dann gilt:

- a) $(u_h, \gamma_h) \geq R_h(u_h, \gamma_h) \geq (u_h^*, \gamma_h^*) \Rightarrow (u_h, \gamma_h) \geq (u_h + v_h, \gamma_h) \geq R_h(u_h + v_h, \gamma_h) \geq (u_h^*, \gamma_h^*),$
- b) $(u_h, \gamma_h) \leq R_h(u_h, \gamma_h) \leq (u_h^*, \gamma_h^*) \Rightarrow (u_h, \gamma_h) \leq (u_h + v_h, \gamma_h) \leq R_h(u_h + v_h, \gamma_h) \leq (u_h^*, \gamma_h^*).$

Das obige Lemma entspricht den Lemmata 3.1 und 3.2 in [25]. Der folgende Beweis unterscheidet sich nur in unwesentlichen Details von dem dortigen.

Beweis:

Wir zeigen a). Für beliebige $P_i \in P_h$ definieren wir einen Relaxationsoperator mit Projektion $R_h^{+(i)} : S_h \times L_h \rightarrow S_h \times L_h$ punktweise durch:

$$(4.4) \quad (R_h^{+(i)}(u_h, \gamma_h))^j := \begin{cases} (u_h^j, \gamma_h^j) & \text{falls } i \neq j, \\ ((R_h^u(u_h, \gamma_h))^j, 1) & \text{falls } i = j. \end{cases}$$

Sei $(u_h, \gamma_h) \in S_h \times L_h$ eine Oberlösung, die der Ungleichung

$(u_h, \gamma_h) \geq (u_h^*, \gamma_h^*)$ genügt. Wir betrachten die Folge $(u_h^{(v)}, \gamma_h^{(v)})_{v \in \mathbb{N}}$ mit

$$(4.5) \quad \begin{cases} (u_h^{(0)}, \gamma_h^{(0)}) := (u_h, \gamma_h), \\ (u_h^{(v+1)}, \gamma_h^{(v+1)}) := G_h^+(u_h^{(v)}, \gamma_h^{(v)}), \quad v \in \mathbb{N}, \end{cases}$$

$$G_h^+ := R_h^{+(i_m)} \circ R_h^{+(i_{m-1})} \circ \dots \circ R_h^{+(i_1)}, D_h \cap N_h(\gamma_h) = \{P_{i_1}, \dots, P_{i_m}\}.$$

Für alle $v \in \mathbb{N}$ gilt $\gamma_h^{(v)} = \gamma_h$.

Aus (3.3) und (4.4) folgert man

$$(u_h, \gamma_h) \geq R_h(u_h, \gamma_h) \geq (u_h^*, \gamma_h^*) \Rightarrow$$

$$(u_h, \gamma_h) \geq R_h^{+(i)}(u_h, \gamma_h) \geq R_h(u_h, \gamma_h) \geq R_h(R_h^{+(i)}(u_h, \gamma_h)) \geq (u_h^*, \gamma_h^*) \quad \forall P_i \in N_h$$

Durch wiederholte Anwendung dieser Implikation erkennt man

$$(4.6) \quad (u_h, \gamma_h) \geq G_h^+(u_h, \gamma_h) \geq R_h(G_h^+(u_h, \gamma_h)) \geq (u_h^*, \gamma_h^*).$$

Aus (4.5) und (4.6) folgt induktiv

$$(4.7) \quad (u_h^{(v)}, \gamma_h^{(v)}) \geq (u_h^{(v+1)}, \gamma_h^{(v+1)}) \geq R_h(u_h^{(v+1)}, \gamma_h^{(v+1)}) \geq (u_h^*, \gamma_h^*) \quad \forall v \in \mathbb{N}.$$

Als monoton fallende und beschränkte Folge besitzt $(u_h^{(v)})_{v \in \mathbb{N}}$ einen

Grenzwert $u_h^\infty \in S_h$. Dabei gilt

$$u_h^\infty(P) = u_h(P) \quad \forall P \in P_h^0 \cup P_h^+ \cup (P_h \setminus N_h(\gamma_h)), \quad u_h^\infty(P) \geq 0 \quad \forall P \in P_h.$$

Aufgrund der Stetigkeit von R_h gilt wegen (4.7) auch

$$(u_h, \gamma_h) \geq (u_h^\infty, \gamma_h) \geq R_h(u_h^\infty, \gamma_h) \geq (u_h^*, \gamma_h^*).$$

Es bleibt zu zeigen, daß $u_h^\infty - u_h$ die (eindeutig bestimmte) Lösung von (4.3)

ist. Aus der Definition (4.4) folgt

$$\begin{aligned}
 (R_h(u_h^\infty, \gamma_h))^i &= (R_h^{(i)}(u_h^\infty, \gamma_h))^i \\
 \begin{cases} = (R_h^{(i)}(u_h^\infty, \gamma_h))^i = (u_h^\infty, \gamma_h)^i & \text{für } p_i \in D_h \cap N_h(\gamma_h) \text{ mit } u_h^{\infty, i} > 0, \\ \leq (R_h^{(i)}(u_h^\infty, \gamma_h))^i = (u_h^\infty, \gamma_h)^i & \text{für } p_i \in D_h \cap N_h(\gamma_h) \text{ mit } u_h^{\infty, i} = 0. \end{cases}
 \end{aligned}$$

Wie im Beweis zu Lemma 3.1 und zur Folgerung 3.4 schließen wir

$$(A_h \underline{u}_h^\infty + E_h \underline{\gamma}_h)^i = 0 \quad \forall p_i \in D_h \cap N_h(\gamma_h) \text{ mit } u_h^{\infty, i} > 0,$$

$$(A_h \underline{u}_h^\infty + E_h \underline{\gamma}_h)^i \geq 0 \quad \forall p_i \in D_h \cap N_h(\gamma_h).$$

□

4.3 Numerische Berechnung des Hilfsproblems

Die näherungsweise Berechnung des Hilfsproblems (4.3) erfolgt mit Hilfe eines Zweigitterverfahrens. Aus Arbeiten von A. Brandt & C.W. Cryer [10] und J. Mandel [23] ist bekannt, wie ein Mehrgitter-Algorithmus zur Lösung von Hindernisproblemen entworfen werden kann. Der wesentliche Unterschied zwischen beiden Ansätzen besteht in der Übertragung der Nebenbedingung auf das grobe Gitter. In [10] wird das Energiefunktional in einem zu großen, in [23] in einem recht kleinen Kegel minimiert. Wir folgen bei unserem Algorithmus der Arbeit von A. Brandt & C.W. Cryer und kontrollieren die Nicht-Negativität von $u_h + v_{2h}$ nur in Grob-Gitter-Punkten. Wie schon der Vergleich der numerischen Ergebnisse in [10] und [23] vermuten läßt, liefert dieses Vorgehen im Standardfall die schnelleren Algorithmen; lediglich in Spezialfällen mit extrem kleinen Hindernissen, die ansonsten auf dem groben Gitter nicht mehr sichtbar sind, ist der in [23] beschriebene Weg auch unter praktischen Gesichtspunkten vorzuziehen [24].

Wir berechnen die Lösung v_h des Hilfsproblems (4.3) nur sehr grob mit Hilfe eines Zweigitter-Zyklus:

Dem Problem (4.3) entspricht das Grob-Gitter-Problem

$$(4.8) \quad \left\{ \begin{array}{l} \text{Gesucht ist } v_{2h} \in \tilde{S}_{2h}(u_h, \gamma_h), \text{ so da\ss} \\ (\psi_{2h} - v_{2h})^T A_{2h} v_{2h} \geq -(\psi_{2h} - v_{2h})^T I_{2h}^T (A_h u_h + E_h \gamma_h) \quad \forall \psi_{2h} \in \tilde{S}_{2h}(u_h, \gamma_h). \end{array} \right.$$

Zur Berechnung von (4.8) f"uhren wir, beginnend mit dem Startwert $v_{2h} \equiv 0$, einige Schritte des Gau\ss-Seidel-Verfahrens mit Projektion nach C.W. Cryer [15] aus. Dann interpolieren wir $v_h := I_{2h}^h v_{2h}$ und gl"atten v_h durch einige weitere Gau\ss-Seidel-Relaxationen mit Projektion auf Ω_h . Wir k"onnen nicht erwarten, da\ss $(u_h + v_h, \gamma_h)$ weiterhin Ober- bzw. Unterl"osung ist. Daher wird als neue N"aherung f"ur den Druck die maximale Konvexkombination zwischen u_h und $u_h + v_h$ gew"ahlt, die die entsprechende Eigenschaft besitzt.

4.4 Der Zweigitter-Algorithmus

Die obigen Betrachtungen f"uhren zu folgendem

Algorithmus MGDAM:

a) Start.

Falls keine geeigneteren Startwerte bekannt sind, setze

$$u_h^- \equiv \gamma_h^- \equiv 0,$$

$$\gamma_h^+ \equiv 1, \quad u_h^+ = K - \underline{x} \underline{e} \quad \text{mit } K := \sup_{\underline{x} \in P_h} (\underline{x} \underline{e} + u_h^0(\underline{x})).$$

b) Relaxation von u_h und γ_h .

F"uhre v_1 mal

$$(u_h^-, \gamma_h^-) \leftarrow G_h(u_h^-, \gamma_h^-),$$

$$(u_h^+, \gamma_h^+) \leftarrow G_h(u_h^+, \gamma_h^+)$$

aus.

c) Berechnung einer Korrektur für u_h in den Punkten der gesättigten Zone.

Bestimme $\underline{b}_h^- \leftarrow -A_h \underline{u}_h^- - E_h \underline{\gamma}_h^-$ und restringiere $\underline{b}_{2h}^- \leftarrow I_{2h}^T \underline{b}_h^-$.

Beginnend mit dem Startwert $\underline{v}_{2h}^- \equiv 0$ führe v_2 mal

$$\underline{v}_{2h}^- \leftarrow \text{PGS}_{2h}(\underline{v}_{2h}^-, \underline{b}_{2h}^-, -\underline{u}_h^-, N_{2h}(\underline{\gamma}_h^-))$$

aus und interpoliere

$$\underline{v}_h^- \leftarrow I_{2h}^h \underline{v}_{2h}^-.$$

Berechne entsprechend eine Grob-Gitter-Korrektur $\underline{v}_h^+$ für die Oberlösung.

Anschließend führe v_3 mal

$$\underline{v}_h^- \leftarrow \text{PGS}_h(\underline{v}_h^-, \underline{b}_h^-, -\underline{u}_h^-, N_h(\underline{\gamma}_h^-)),$$

$$\underline{v}_h^+ \leftarrow \text{PGS}_h(\underline{v}_h^+, \underline{b}_h^+, -\underline{u}_h^+, N_h(\underline{\gamma}_h^+))$$

aus.

d) Addition der berechneten Korrekturen.

Setze $\underline{u}_h^- \leftarrow \underline{u}_h^- + \lambda \underline{v}_h^-$, wobei $\lambda \in [0,1]$ maximal gewählt wird, so daß

$(\underline{u}_h^-, \underline{\gamma}_h^-)$ Unterlösung bleibt. Setze entsprechend $\underline{u}_h^+ \leftarrow \underline{u}_h^+ + \rho \underline{v}_h^+$.

e) Ende-Abfrage.*

Falls $\|\underline{u}_h^+ + \underline{\gamma}_h^+ - \underline{u}_h^- - \underline{\gamma}_h^-\|_\infty \leq \epsilon$ gilt, beende den Algorithmus, andernfalls gehe nach Schritt b).

Dabei sei der Gauß-Seidel-Operator mit Projektion $\text{PGS}_H(w_H, b_H, \psi_H, N_H)$ wie folgt erklärt:

$$(4.9) \quad \begin{cases} \tilde{w}_H^i := \begin{cases} w_H^i & \text{falls } p_i \notin N_H \cap D_H \\ \max\{\psi_H(p_i), (b_H^i - \sum_{j=1}^{i-1} a_{ij}^H \tilde{w}_H^j - \sum_{j=i+1}^n a_{ij}^H w_H^j) / a_{ii}^H\} & \text{falls } p_i \in N_H \cap D_H, \end{cases} \\ \text{PGS}_H(w_H, b_H, \psi_H, N_H) := \tilde{w}_H. \end{cases}$$

*Diese Ende-Abfrage ist zunächst nur bei Problemen mit eindeutiger Lösung sinnvoll. Die Behandlung von Problemen mit mehrdeutiger Lösung wird anhand eines Beispiels in Kapitel 6 erläutert.

Bemerkung 4.2

a) Wegen $N_{2h}(\gamma_h) \subseteq N_h(\gamma_h)$ gilt für die in Schritt c) berechneten Korrekturen

$$v_h^i = 0 \quad \forall p_i \in (P_h \setminus N_h(\gamma_h)) \cup P_h^0 \cup P_h^+.$$

Nach jedem Iterationsschritt von MGDAM sind $u_h^-, u_h^+ \in M_h(u_h^0)$, und beide Näherungen genügen der Verträglichkeitsbedingung (1.2a).

b) Es gilt stets $v_h^- \geq 0$ und $v_h^+ \leq 0$.

Die erste Ungleichung deutet an, daß die Nebenbedingung bei der Minimierung des Energiefunktionalis der Unterlösung in (4.8) überflüssig ist.

Wir zeigen $v_h^- \geq 0$.

Wie im Beweis von Lemma 3.2 und Folgerung 3.4 erkennt man

$$b_h^{-i} = (-A_h u_h^- - E_h \gamma_h^-)^i \geq 0 \quad \forall p_i \in D_h$$

und daher

$$b_{2h}^{-i} = (I_{2h}^T b_h)^i \geq 0 \quad \forall p_i \in D_{2h}.$$

Aus (4.9) folgt $v_{2h}^- \geq 0$, also auch $v_h^- = I_{2h}^h v_{2h}^- \geq 0$. ┐

Bei entsprechenden Werten für γ_h ist H.W. Alt's Relaxationsoperator R_h invers monoton bzgl. u_h :

Lemma 4.3

Seien $(u_h^-, \gamma_h^-), (u_h^+, \gamma_h^+) \in S_h \times L_h$ und gelte $u_h^+(P) \geq 0, \gamma_h^-(P) \leq \gamma_h^+(P) \leq 1$ für alle $P \in P_h^0 \cup P_h^+$ sowie $u_h^-(P) \leq u_h^+(P)$ für alle $P \in P_h^0 \cup P_h^+$. Außerdem genüge (u_h^-, γ_h^-) der Verträglichkeitsbedingung (1.2a). Dann folgt aus

$$\left. \begin{array}{l} (A_h u_h^+ + E_h \gamma_h^+)^i \geq 0, \\ u_h^{-i} (A_h u_h^- + E_h \gamma_h^-)^i \leq 0 \end{array} \right\} \quad \forall p_i \in D_h$$

die Beziehung $u_h^- \leq u_h^+$.

Beweis:

Sei $(u_h^+ - u_h^-)_-^i := -\min\{0, (u_h^+ - u_h^-)^i\}$ und $(u_h^+ - u_h^-)_+^i := (u_h^+ - u_h^-)^i + (u_h^+ - u_h^-)_-^i$.

Wegen $(u_h^+ - u_h^-)_- [A_h(u_h^+ - u_h^-) + E_h(\gamma_h^+ - \gamma_h^-)] \geq 0$ gilt auch

$$\begin{aligned} (u_h^+ - u_h^-)_- A_h(u_h^+ - u_h^-)_- &\leq (u_h^+ - u_h^-)_- [A_h(u_h^+ - u_h^-)_+ + E_h(\gamma_h^+ - \gamma_h^-)] \\ &\leq \sum_{i=1}^n (u_h^+ - u_h^-)_-^i [a_{ii}(u_h^+ - u_h^-)_+^i + e_{ii}(\gamma_h^+ - \gamma_h^-)^i] \\ &\leq \sum_{i=1}^n (u_h^+ - u_h^-)_-^i e_{ii}(\gamma_h^+ - \gamma_h^-)^i; \end{aligned}$$

und aus der Verträglichkeitsbedingung (1.2a) für (u_h^-, γ_h^-) folgt

$$(u_h^+ - u_h^-)_-^i \neq 0 \Rightarrow u_h^{-i} > u_h^{+i} \geq 0 \Rightarrow \gamma_h^{-i} = 1 \geq \gamma_h^{+i} \Rightarrow (\gamma_h^+ - \gamma_h^-)^i = 0. \quad \square$$

Infolge von Lemma 3.1, Lemma 3.2, Bemerkung 4.2 und Lemma 4.3 gilt:

Folgerung 4.4

Die durch MGDAM definierte Iterationsfolge (u_h^+, γ_h^+) konvergiert gegen den größten Fixpunkt von R_h unterhalb des Startwertes für die Oberlösung; entsprechend konvergiert (u_h^-, γ_h^-) gegen den kleinsten Fixpunkt von R_h oberhalb des Startwertes für die Unterlösung.

4.5 Automatische Steuerung der Anzahl der Relaxationen

Die optimalen Werte für v_1 und v_2 sind problemabhängig. Will man ein universell anwendbares Programm erstellen, so empfiehlt es sich, das Mehrgitterverfahren in Abhängigkeit von der jeweiligen Situation selbst entscheiden zu lassen, wie häufig relaxiert werden soll. Anstelle einer festen Anzahl von Relaxationsschritten auf jeder Stufe werden zwei Steuerparameter n und δ vorgegeben. Bei jeder Relaxation in Schritt b) wird die ℓ_2 -Norm der Änderung Δ in u_h bestimmt. Bezeichne Δ^{alt} den entsprechenden Wert der vorangegangenen Relaxation. Solange $\Delta \leq n \cdot \Delta^{\text{alt}}$ gilt, ist die Konvergenzgeschwindigkeit der Relaxation noch hinreichend gut, und es lohnt ein weiterer Glättungsschritt. Ansonsten wird gemäß Schritt c)

eine Grob-Gitter-Korrektur für u_h berechnet. Als Toleranz für die Berechnung dieser Korrektur wird $\text{Tol} = \delta \cdot \Delta$ gewählt, d.h. in Schritt c) wird v_{2h}^- bzw. v_{2h}^+ solange relaxiert, bis die ℓ_2 -Norm der Änderung kleiner als Tol ist. Verfahren mit einer derartigen automatischen Steuerung ("accommodative algorithms") wurden von A. Brandt entwickelt ([7],[8],[9],[10]).

4.6 Die Full Multigrid Version

Als Startwerte für den Zweigitter-Algorithmus können jede Oberlösung (u_h^+, γ_h^+) und jede Unterlösung (u_h^-, γ_h^-) (in dem in Abschnitt 3.1 vereinbarten Sinne) mit $(u_h^+, \gamma_h^+) \geq (u_h^-, \gamma_h^-)$ verwendet werden. In der Praxis erweist sich die Benutzung der "Full Multigrid" Technik, auch "nested iteration" genannt, als entscheidender Vorteil. Die Iteration beginnt auf Ω_{h_0} mit den bekannten Startwerten für (u_h^+, γ_h^+) und (u_h^-, γ_h^-) . In aufsteigender Reihenfolge werden auf jedem Gitter Ω_{h_k} Ober- wie Unterlösung solange mit Hilfe eines Relaxations- bzw. Mehrgitterverfahrens iteriert, bis $\|u_h^+ + \gamma_h^+ - u_h^- - \gamma_h^-\|_\infty$ kleiner als eine vorgegebene Toleranz ϵ_k ist. Dann erfolgt die Interpolation von (u_h^-, γ_h^-) und (u_h^+, γ_h^+) auf das nächst feinere Gitter. Wir betonen, daß $(u_{2h}, \gamma_{2h}) \leq R_{2h}(u_{2h}, \gamma_{2h})$ nicht $(u_{2h}, \gamma_{2h}) \leq R_h(u_{2h}, \gamma_{2h})$ impliziert. Falls jedoch ϵ nicht zu klein gewählt wurde, wird eine Ω_{2h} -Ober- bzw. Ω_{2h} -Unterlösung durch wenige Iterationen mit G_h in eine Ω_h -Ober- bzw. Ω_h -Unterlösung überführt.

4.7 Bemerkung zu anderen Varianten des Algorithmus

Es liegt der Gedanke nahe, in Schritt c) die a-posteriori-Glättung der berechneten Korrektur für u_h^- bzw. u_h^+ durch Relaxationsschritte für u_h und γ_h gemäß (3.1) bzw. (3.5) zu ersetzen. Dieses Vorgehen führt jedoch dazu,

daß nur kleine Anteile der berechneten Korrekturen aufaddiert werden können, ohne die Ober- bzw. Unterlösungseigenschaft zu verletzen.

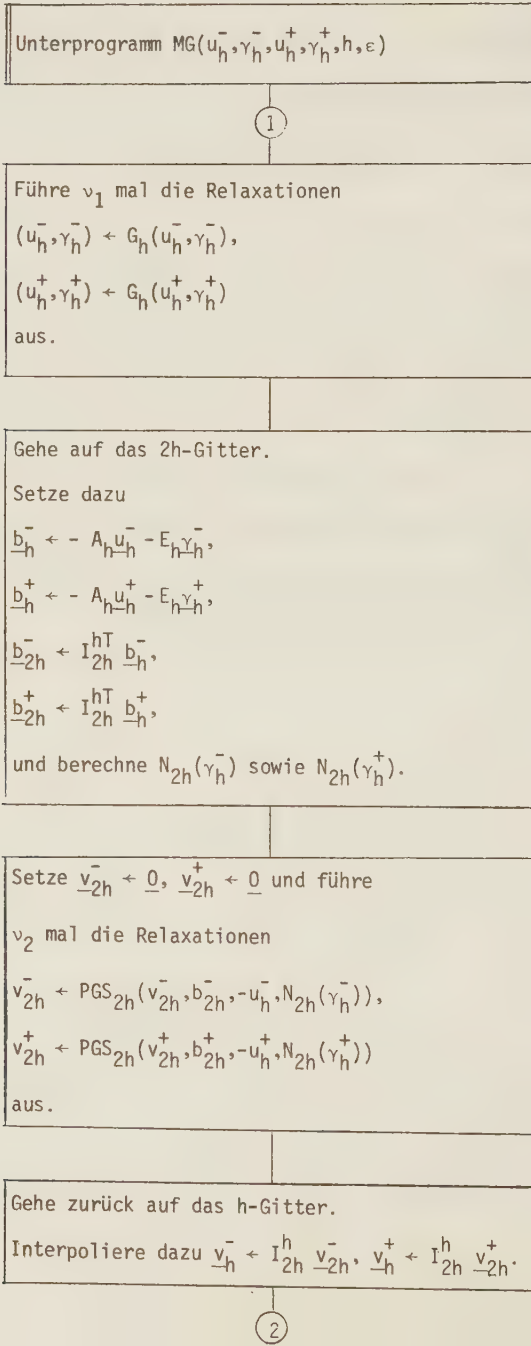
Derselbe Effekt tritt ein, wenn anstelle des Hilfsproblems (4.3) die entsprechende Poissongleichung ohne Nebenbedingung gewählt wird.

Bei der in Abschnitt 4.4 beschriebenen Version von MGDAM wird auch durch Addition der vollen in Schritt c) berechneten Korrektur die Konvergenzgeschwindigkeit nicht wesentlich vergrößert. Der Konvergenzbeweis verliert bei diesem Vorgehen jedoch seine Gültigkeit.

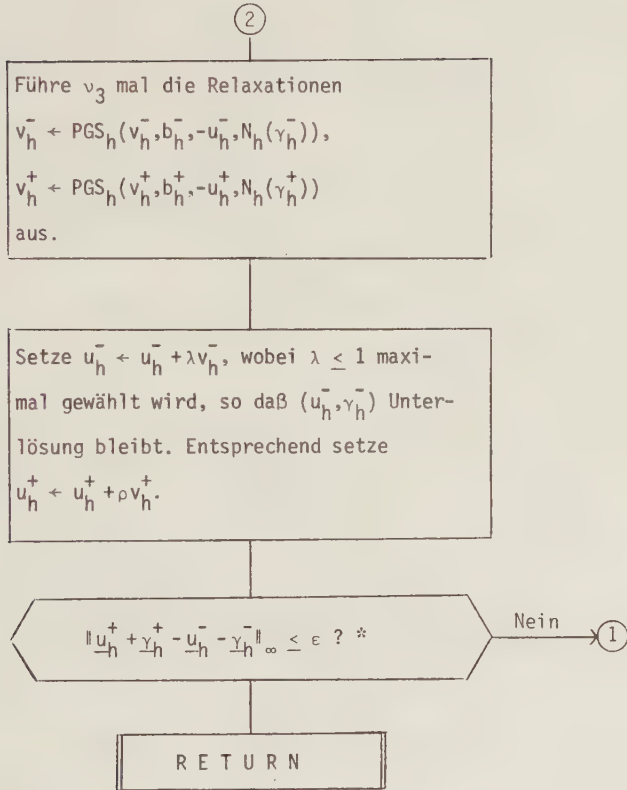
Unserer Erfahrung nach führt die Berechnung von (4.3) mit Hilfe von 3-Gitter-Algorithmen zu langsameren Verfahren. Für wachsende Gitterzahl wird der Abstand der Menge der aktiven Punkte $N_h(\gamma_h)$ zum freien Rand offenbar zu groß.

Durch zusätzliche Relaxationen in der Nähe des freien Randes läßt sich die Konvergenzgeschwindigkeit der in Abschnitt 4.4 beschriebenen Version von MGDAM leicht verbessern. (Die Verbesserung beträgt bei dem in Kapitel 6 beschriebenen Testproblem 0 ca 5%).

Flußdiagramm 1: MGDAM ohne automatische Steuerung und ohne Startwertzuweisung

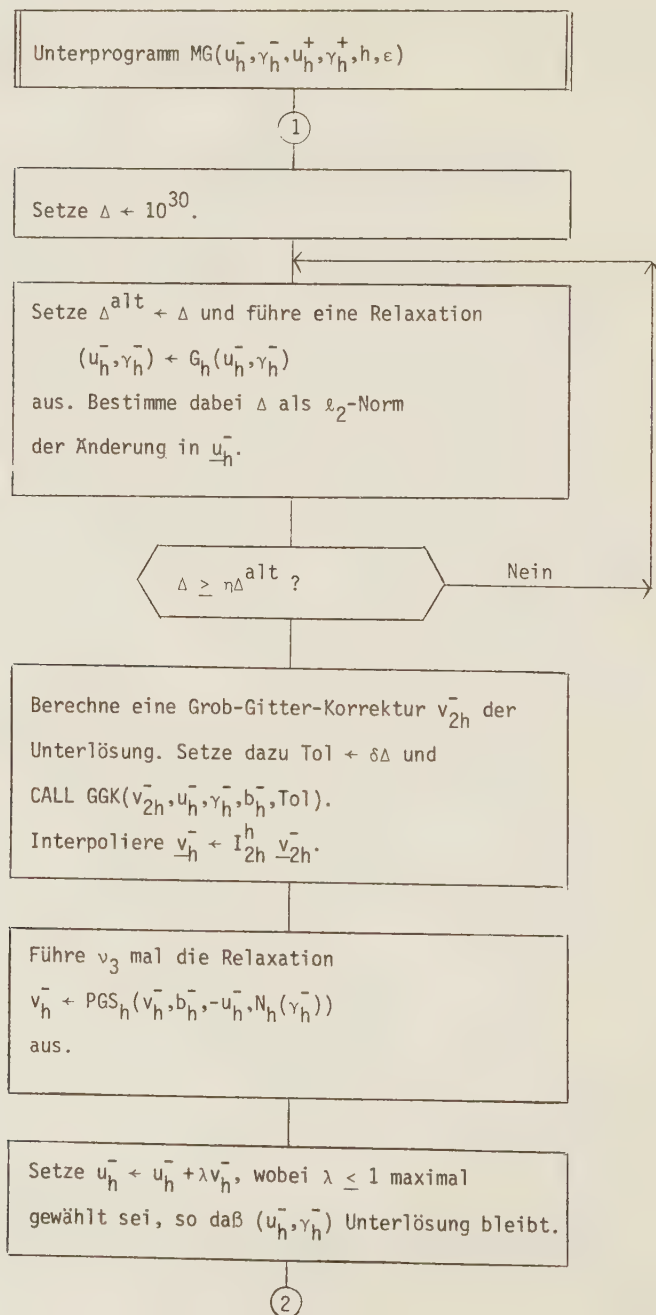


Fortsetzung Flußdiagramm 1

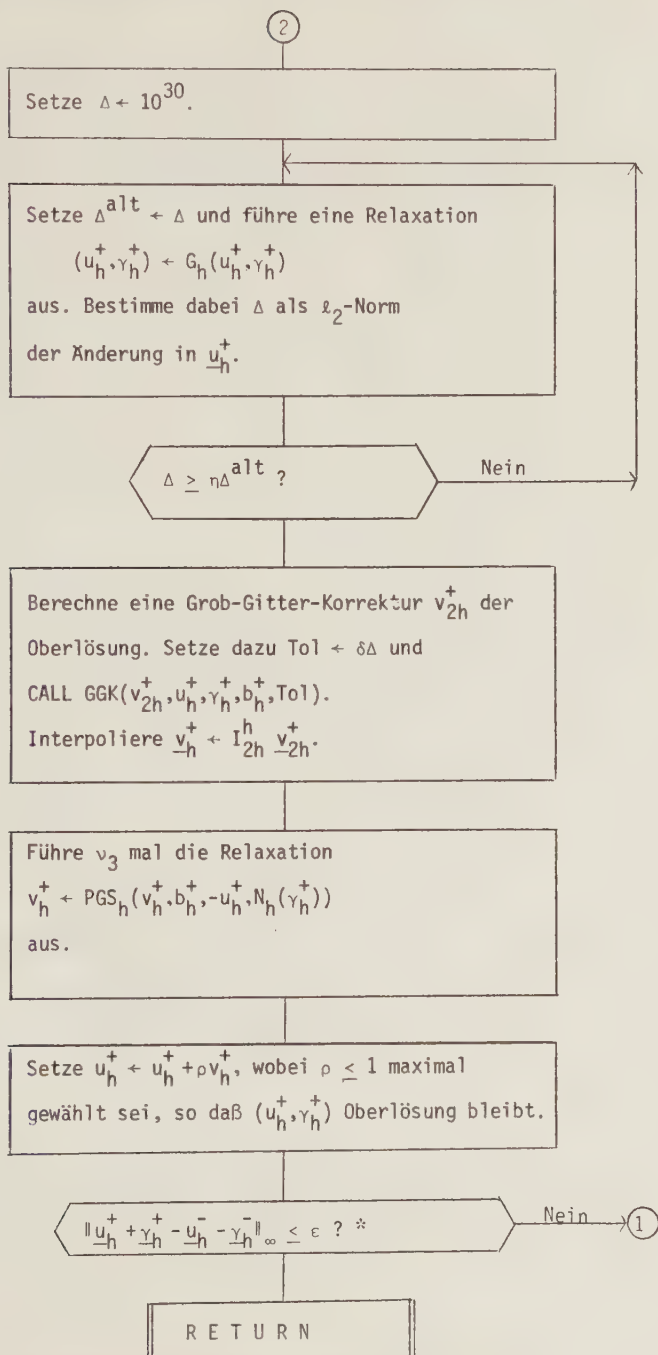


* Zur Handhabung der Ende-Abfrage beachte man die Fußnote auf Seite 31.

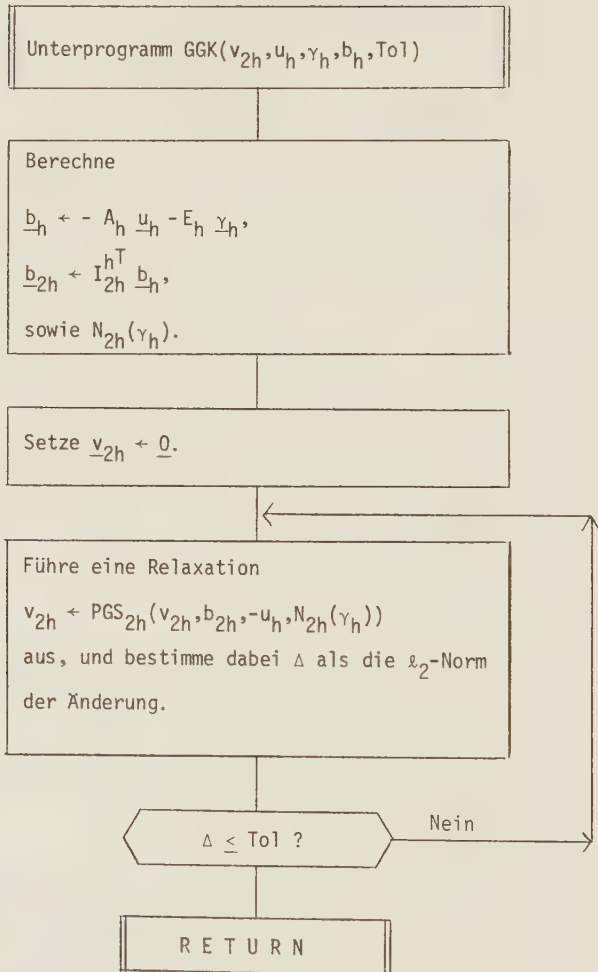
Flußdiagramm 2: MGDAM mit automatischer Steuerung und ohne Startwertzuweisung



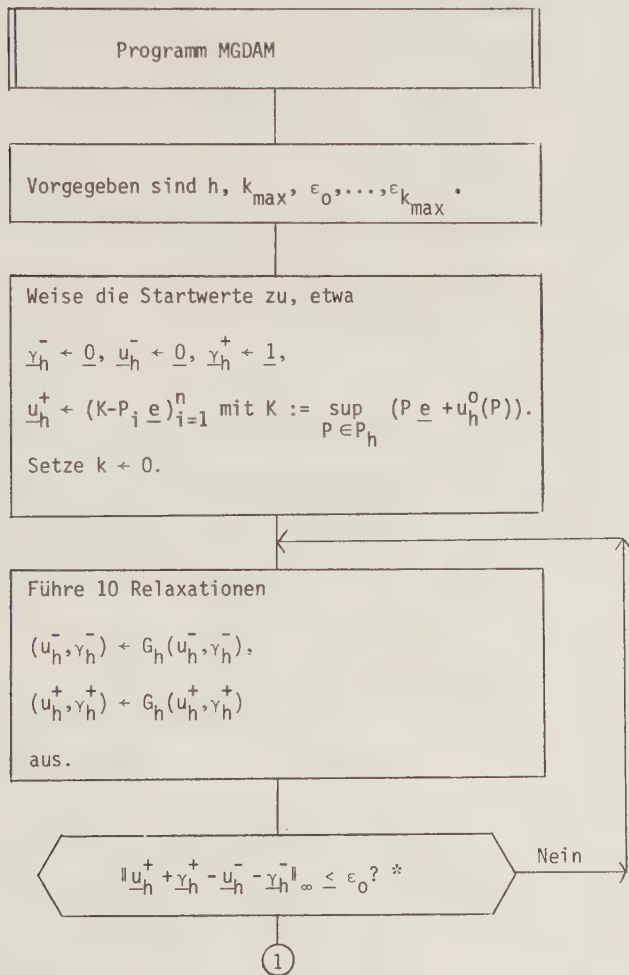
Fortsetzung Flußdiagramm 2



Fortsetzung Flußdiagramm 2

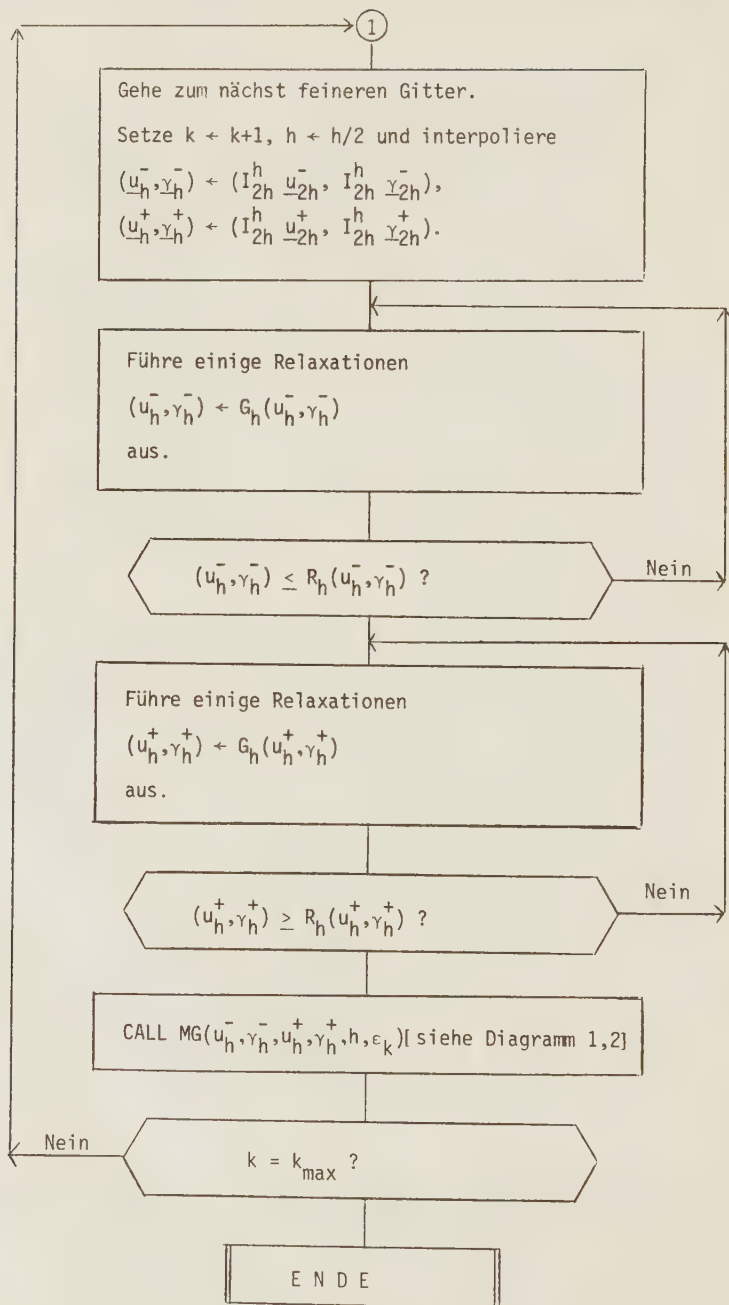


Flußdiagramm 3: Die Full Multigrid Version von MGDAM



* Zur Handhabung dieser Abfrage beachte man die Fußnote auf Seite 31.

Fortsetzung von Flußdiagramm 3



5. Der Algorithmus FASDAM mit voller Grob-Gitter-Korrektur

Charakteristisch für Mehrgitterverfahren ist die Tatsache, daß der Fehler pro Iterationsschritt um einen von der Schrittweite unabhängigen Faktor reduziert wird. Dieses Verhalten zeigt der Fehler bei Anwendung des in Kapitel 4 vorgestellten Verfahrens nicht. Der Grund dafür ist, wie numerische Experimente zeigten, nicht etwa die Dämpfung der Grob-Gitter-Korrektur in Schritt d), sondern die schlechte Approximation der γ -Werte. Daher entwickelten wir den Algorithmus FASDAM, der im Sinne des "full approximation scheme" (FAS) eine Grob-Gitter-Näherung für u_h und γ_h vorsieht. Die Konstruktion der Grob-Gitter-Korrektur bedarf einiger Sorgfalt.

5.1 Bemerkungen zur allgemeinen Konstruktion von Grob-Gitter-Korrekturen

Bei der Diskretisierung linearer elliptischer Probleme wie der Poisson-Gleichung mit Hilfe der Methode der Finiten Elemente entstehen lineare Gleichungssysteme der Gestalt $A_h u_h = f_h$. Das zugehörige Grob-Gitter-Problem

$$A_{2h} v_{2h} = I_h^{2h} (f_h - A_h u_h) \text{ mit } A_{2h} = I_{2h}^{hT} A_h I_{2h}^h, \quad I_h^{2h} = I_{2h}^{hT}$$

ist dort natürlich gegeben.

Will man sich in der Praxis nicht auf den natürlichen "residual weighting" Operator $I_h^{2h} = I_{2h}^{hT}$ festlegen lassen, so bietet sich der besonders leicht zu berechnende Injektionsoperator (Identität in allen Grob-Gitter-Punkten) mit entsprechender Wichtung an. Dieser Operator besitzt außerdem den Vorteil, daß der Defekt $f_h - A_h u_h$ nur in den Punkten des groben Gitters berechnet werden muß, und liefert bei der Lösung einfacher elliptischer Probleme mit glatten Koeffizienten optimale Mehrgitterverfahren ([7]; [9], 4.4; [27], 3.6.1). Es ist jedoch bekannt, daß die Injektion weniger robust

ist als eine Restriktion mit Wichtung über mehrere Punkte ([9], 4.6) und nur bei glattem Defekt verwendet werden darf ([27], 3.6.1).

Die Mehrgitter Technik kann auch zur direkten Behandlung von Problemen der Gestalt $F_h(\underline{u}_h) = \underline{0}$ mit nichtlinearem (diskreten) Differentialoperator F_h eingesetzt werden. In diesem Fall wird auf dem groben Gitter eine Approximation für die volle Lösung, und nicht nur eine Korrektur, als (näherungsweise) Lösung des Problems

$$(5.1) \quad F_{2h}(\underline{u}_{2h}) = F_{2h}(\tilde{I}_h^{2h} \underline{u}_h) - I_h^{2h} F_h(\underline{u}_h)$$

bestimmt. Dabei bezeichnen I_h^{2h} , $\tilde{I}_h^{2h}$ zwei Restriktionsoperatoren.

Die aktuelle Näherung $\underline{u}_h$ wird anschließend mit $I_{2h}^h(\underline{u}_{2h} - \tilde{I}_h^{2h} \underline{u}_h)$ korrigiert.

Bei einem iterativen Verfahren zur Lösung von (5.1) sollte unbedingt

$\tilde{I}_h^{2h} \underline{u}_h$ als Startwert verwendet werden. (Siehe [19], 7 ; [9], 8.)

Die Terme auf der rechten Seite von (5.1) sollen gewährleisten, daß die

Restriktion $\tilde{I}_h^{2h} \underline{u}_h^*$ einer Fein-Gitter-Lösung $\underline{u}_h^*$ das Grob-Gitter-Problem

(5.1) löst und somit $\underline{u}_h^*$ stationärer Punkt des Mehrgitterverfahrens ist.

Eine Anwendung dieses "full approximation scheme" (FAS) auf Hindernisprobleme findet man in [10].

5.2 Die Grob-Gitter-Korrektur bei der Berechnung des diskreten

Dammproblems

Bei konsequenter Anwendung der FAS Technik auf das diskrete Dammproblem mit rechter Seite $\underline{b}_h$ erhalten wir folgendes Grob-Gitter-Problem:

$$(5.2) \quad \left\{ \begin{array}{l} \text{Gesucht sind } \underline{u}_{2h} \in M_{2h}(\underline{u}_{2h}^0) \text{ und } \gamma_{2h} \in L_{2h}, \text{ so daß in allen} \\ \text{Punkten aus } P_{2h} \text{ die Verträglichkeitsbedingung (1.2a) erfüllt} \\ \text{ist und} \\ (A_{2h} \underline{u}_{2h} + E_{2h} \gamma_{2h})^j \quad \left\{ \begin{array}{ll} = b_{2h}^j & \text{falls } p_j \in D_{2h}, \\ \leq b_{2h}^j & \text{falls } p_j \in P_{2h}^0. \end{array} \right. \end{array} \right.$$

Dabei sei

$$\begin{aligned} \underline{b}_{2h} &:= A_{2h} \tilde{I}_h^{2h} \underline{u}_h + E_{2h} \tilde{I}_h^{2h} \underline{y}_h - I_h^{2h} (A_h \underline{u}_h + E_h \underline{y}_h - \underline{b}_h), \\ \underline{u}_{2h}^0 &:= \tilde{I}_h^{2h} \underline{u}_h^0. \end{aligned}$$

Da im Regelfall die spezielle Wahl des Restriktionsoperators $\tilde{I}_h^{2h}$ ("FAS solution averaging") unerheblich für die Konvergenzgeschwindigkeit von Mehrgitter-Algorithmen ist ([8], 3.4), haben wir sowohl für die Restriktion von $\underline{u}_h$ als auch für die Restriktion von $\underline{y}_h$ den einfachsten Operator, die Injektion gewählt.

Bei der Konstruktion von Grob-Gitter-Problemen für Variationsungleichungen langt der auf der rechten Seite auftretende Shift $F_{2h}(\tilde{I}_h^{2h} \underline{u}_h) - I_h^{2h} F_h(\underline{u}_h)$ i.a. nicht aus, um die stationären Punkte zu erhalten. Zusätzlich müssen die Komplementaritätsbedingungen in der richtigen Weise transferiert werden. Die sich bei der Lösung von Hindernisproblemen daraus ergebenden Forderungen an den Restriktionsoperator I_h^{2h} wurden in [10] genau analysiert (siehe Lemma 5.1). Beim diskreten Dammproblem ist die Situation anders, Komplementaritätsbedingungen treten nur in Randpunkten auf. (Siehe Folgerung 3.3.) Bei trivialer rechter Seite $\underline{b}_h = \underline{0}$ treten diese Bedingungen sogar nur in den Punkten aus P_h^0 auf (Folgerung 3.4). Wir wollen jedoch auch die auf gröberen Gittern auftretenden Probleme betrachten und daher $\underline{b}_h \neq \underline{0}$ zulassen.

Um die Anforderungen an den "residual weighting" Operator I_h^{2h} in unserem Fall genauer zu studieren, sei (u_h^*, γ_h^*) Fixpunkt von R_h und $(u_{2h}^*, \gamma_{2h}^*) = (\tilde{I}_h^{2h} u_h^*, \tilde{I}_h^{2h} \gamma_h^*)$ Startwert für die Iteration auf dem $2h$ -Gitter. Sinnvollerweise sollte die Grob-Gitter-Korrektur für (u_h^*, γ_h^*) Null betragen, also $(u_{2h}^*, \gamma_{2h}^*)$ Fixpunkt von R_{2h} sein. Gemäß Folgerung 3.3 muß dann

$$(A_{2h-2h} u_{2h}^* + E_{2h} \gamma_{2h}^* - b_{2h})^i u_{2h}^{*i} = 0 \quad \text{für alle } P_i \in D_{2h} \text{ mit } e_{ii}^{2h} = 0$$

gelten. Daraus folgt aber

$$(I_h^{2h} (A_{h-h} u_h^* + E_h \gamma_h^* - b_h))^i u_h^{*i} = 0 \quad \text{für alle } P_i \in D_{2h} \text{ mit } e_{ii}^{2h} = 0.$$

Entsprechend muß für $P_i \in P_{2h}^0$ mit $e_{ii}^{2h} \neq 0$ die Gleichung

$$(I_h^{2h} (A_{h-h} u_h^* + E_h \gamma_h^* - b_h))^i (1 - \gamma_h^i) = 0$$

erfüllt sein. Kurz gesprochen, es soll

$$\begin{aligned} \forall P_i \in D_{2h} \text{ mit } e_{ii}^{2h} = 0 \quad d_h^i u_h^i = 0 &\Rightarrow (I_h^{2h} d_h)^i u_h^i = 0, \\ \forall P_i \in P_{2h}^0 \text{ mit } e_{ii}^{2h} \neq 0 \quad d_h^i (1 - \gamma_h^i) = 0 &\Rightarrow (I_h^{2h} d_h)^i (1 - \gamma_h^i) = 0 \end{aligned}$$

gelten. Aufgrund der Mittelungseigenschaft führt der Gebrauch des natürlichen "residual weighting" Operators I_{2h}^{hT} (auch in der Praxis) zu divergenten Algorithmen. Als einfachster Ausweg bietet sich an, den natürlichen Restriktionsoperator in den Randknoten durch die Injektion zu ersetzen:

$$(5.4) \quad (I_h^{2h} d_h)^i := \begin{cases} (I_{2h}^{hT} d_h)^i & \text{falls } P_i \in P_{2h} \cap \Omega, \\ 3d_h^i & \text{sonst.} \end{cases}$$

Für $d_h \equiv 1$ nimmt die natürliche Restriktion $I_{2h}^{hT} d_h$ bei der in Kapitel 6 beschriebenen Triangulierung von Ω auf geraden Randstücken abwechselnd die Werte 3.5 und 2.5 an. Bei praktischen Experimenten erwies sich ein einheitlicher Wichtungsfaktor von 3 in der Definition (5.4) als günstig.

Wählt man in allen Gitterpunkten die weniger stabile Injektion, so erhält man Algorithmen, die in der Praxis zumindest 10% langsamer konvergieren.

Bei einer Mehrgitter-Variante traten sogar Fälle von Divergenz auf.

J. Mandel hat bei der Lösung von Hindernisproblemen ähnliche Phänomene beobachtet ([23], Kapitel 6). Wir erklären uns das Versagen in unserem Fall wie folgt: Für die Lösung des kontinuierlichen Problems (1.1) mit

homogenem isotropen Damm gilt $\gamma = \chi(\{u > 0\})$ (Theorem 4.6 in [4] , Theorem 4.12 in [12]). Also kann nicht erwartet werden, daß bei einem Fixpunkt (u_h^*, γ_h^*) von R_h die zweite Komponente γ_h^* in irgendeiner Form glatt ist. Von daher ist verständlich, daß auch nach mehreren Relaxationen unterhalb des freien Randes stark oszillierende Defekte auftreten. In diesem Fall ist eine Restriktion mit Mittelung über mehrere Punkte angebracht.

Die Modifikation des natürlichen Operators I_{2h}^T in (5.4) ist genügend groß, so daß das Verfahren bei einem Fixpunkt von R_h stationär wird:

Lemma 5.1

Sei (u_h^*, γ_h^*) Fixpunkt von R_h , $\gamma_h^* \geq 0$ und $\underline{u}_{2h}^* := \tilde{I}_h^{2h} u_h^*$, $\underline{\gamma}_{2h}^* := \tilde{I}_h^{2h} \gamma_h^*$, wobei für den Restriktionsoperator $\tilde{I}_h^{2h}$ die Injektion gewählt wurde.

Dann gilt $(u_{2h}^*, \gamma_{2h}^*) = R_{2h}(u_{2h}^*, \gamma_{2h}^*)$.

Beweis:

Als Fixpunkt von R_h mit nichtnegativer zweiter Komponente γ_h^* genügt (u_h^*, γ_h^*) der Verträglichkeitsbedingung (1.2a). Daher lassen sich mit den Bezeichnungen

$$\underline{d}_h := \underline{b}_h - A_h u_h^* - E_h \gamma_h^*, \quad \underline{d}_{2h} := I_h^{2h} \underline{d}_h$$

aufgrund der Folgerung 3.3 folgende Aussagen treffen:

In jedem Punkt $P_i \in D_h$ mit $d_h^i \neq 0$ gilt $d_h^i < 0$ und $u_h^{*i} = e_{ii} = 0$, also ist dann $P_i \in \Gamma$ und $\gamma_h^{*i} = 1$.

Für $P_i \in P_h^0$ gilt $u_h^{*i} = 0$ und

$$d_h^i < 0 \Rightarrow e_{ii} = 0, \quad \gamma_h^{*i} = 1,$$

$$d_h^i > 0 \Rightarrow \gamma_h^{*i} = 1.$$

Für $P_i \in P_h^+$ gilt $\gamma_h^{*i} = 1$ und $u_h^{*i} = u_h^{0,i}$.

Wegen der Definition (5.4) und aufgrund der Beziehung $e_{ii}^{2h} = 0 \Rightarrow e_{ii}^h = 0$ folgt:

Jeder Punkt $P_i \in D_{2h}$ mit $d_{2h}^i \neq 0$ besitzt einen Ω_h -Nachbarn aus $P_h \cap \Gamma$.

Der Punkt P_i ist daher selber aus $P_{2h} \cap \Gamma$, und es gilt

$$d_{2h}^i < 0, u_{2h}^{*i} = e_{ii}^{2h} = 0, \gamma_{2h}^{*i} = 1.$$

Für $P_i \in P_{2h}^0$ ist $u_{2h}^{*i} = 0$ und

$$d_{2h}^i < 0 \Rightarrow e_{ii}^{2h} = 0, \gamma_{2h}^{*i} = 1,$$

$$d_{2h}^i > 0 \Rightarrow \gamma_{2h}^{*i} = 1.$$

Für $P_i \in P_{2h}^+$ gilt $\gamma_{2h}^{*i} = 1$ und $u_{2h}^{*i} = u_{2h}^{0,i}$.

Der Relaxationsoperator R_{2h} läßt sich mit Hilfe von d_{2h} wie folgt beschreiben:

$$R_{2h}^u(u_{2h}^*, \gamma_{2h}^*) = \begin{cases} \max\{(d_{2h}^i + a_{ii}^{2h} u_{2h}^{*i} + e_{ii}^{2h} \gamma_{2h}^{*i} - e_{ii}^{2h})/a_{ii}^{2h}, 0\} & \text{falls } P_i \in D_{2h}, \\ u_{2h}^{0,i} & \text{sonst,} \end{cases}$$

$$R_{2h}^\gamma(u_{2h}^*, \gamma_{2h}^*) = \begin{cases} \min\{(d_{2h}^i + a_{ii}^{2h} u_{2h}^{*i} + e_{ii}^{2h} \gamma_{2h}^{*i})/e_{ii}^{2h}, 1\} & \text{falls } P_i \in D_{2h} \cup P_{2h}^0 \\ & \text{mit } e_{ii}^{2h} \neq 0, \\ 1 & \text{sonst.} \end{cases}$$

Da $(u_{2h}^*, \gamma_{2h}^*)$ der Verträglichkeitsbedingung (1.2a) genügt, folgt die Behauptung. $\square$

5.3 Der Mehrgitter-Algorithmus FASDAM

In den obigen Betrachtungen wurde die Gestalt eines einfachen Zweigitterzyklus des Algorithmus FASDAM weitgehend festgelegt:

Nach einigen Iterationen $(u_h, \gamma_h) \leftarrow G_h(u_h, \gamma_h)$, wobei G_h gemäß (3.6) de-

finiert sein soll, wird mit Hilfe des Grob-Gitter-Problems (5.2) der Operator G_{2h} gebildet. Mit der Injektion der aktuellen Näherung (u_h, γ_h) als Startwert werden einige Relaxationen $(u_{2h}, \gamma_{2h}) \leftarrow G_{2h}(u_{2h}, \gamma_{2h})$ durchgeführt. Anschließend werden die berechneten Korrekturen auf das h-Gitter interpoliert und soweit sinnvoll aufaddiert:

$$\begin{aligned} u_h &\leftarrow \max\{0, u_h + I_{2h}^h(u_{2h} - \tilde{I}_h^{2h} u_h)\}, \\ \gamma_h &\leftarrow \max\{0, \min\{1, \gamma_h + I_{2h}^h(\gamma_{2h} - \tilde{I}_h^{2h} \gamma_h)\}\}. \end{aligned}$$

Es schließen sich gegebenenfalls einige Glättungsschritte auf dem h-Gitter an.

Die Existenz einer Lösung des Grob-Gitter-Problems ist nicht gesichert ([3], Remark 2.9). Für unsere Zwecke ist jedoch allein wichtig, daß durch die Grob-Gitter-Korrektur die langsam schwingenden Komponenten des Fehlers besser gedämpft werden können als durch weitere Iteration mit G_h .

Das obige Zweigitterverfahren kann in der üblichen Weise zu einem echten Mehrgitterverfahren verallgemeinert werden. Auf den gröberen Gittern Ω_{h_k} , $k < k_{\max}$ muß bei der Bestimmung des Defektes wie bei der Bildung von G_{h_k} dann die nichttriviale rechte Seite b_{h_k} berücksichtigt werden.

Ähnlich wie in Kapitel 4 empfiehlt es sich auch hier, das Verfahren selbst entscheiden zu lassen, wie oft relaxiert werden soll und wieviele Grob-Gitter-Korrekturen ausgeführt werden sollen. Das folgende Programm beschreibt in Abhängigkeit von den Steuerparametern η und δ einen Zyklus des Mehrgitter-Algorithmus mit voller Grob-Gitter-Korrektur FASDAM. Dazu bezeichne (u_h, γ_h) eine bereits berechnete Approximation der Lösung.

Algorithmus FASDAM (Ein Iterationsschritt des Algorithmus FASDAM mit automatischer Steuerung)

a) Initialisierung.

Setze $k \leftarrow k_{\max}$, $h \leftarrow h_k$, $\Delta \leftarrow 10^{30}$ und $b_h \leftarrow 0$.

b) A-priori-Glättung.

Setze $\Delta^{\text{alt}} \leftarrow \Delta$ und führe die Relaxation

$$(u_h, \gamma_h) \leftarrow G_h(u_h, \gamma_h)$$

aus. Bestimme dabei Δ als die ℓ_2 -Norm der Änderung in (u_h, γ_h) .

Falls $\Delta < \eta \Delta^{\text{alt}}$ gilt, wiederhole Schritt b). Andernfalls fahre fort mit c).

c) Wechsel zum nächst gröberen Gitter.

Bestimme die Startwerte für die Iteration auf dem nächst gröberen Gitter

$$\underline{u}_{2h} \leftarrow \tilde{I}_h^{2h} \underline{u}_h, \quad \underline{\gamma}_{2h} \leftarrow \tilde{I}_h^{2h} \underline{\gamma}_h;$$

und berechne

$$\underline{b}_{2h} \leftarrow \tilde{I}_h^{2h} (\underline{b}_h - A_h \underline{u}_h - E_h \underline{\gamma}_h) + A_{2h} \underline{u}_{2h} + E_{2h} \underline{\gamma}_{2h}.$$

Setze $k \leftarrow k-1$, $h \leftarrow 2h$.

Bestimme die angestrebte Toleranz ϵ_k als

$$\epsilon_k \leftarrow \delta \Delta$$

und setze $\Delta \leftarrow 10^{30}$.

d) Relaxationsschritt.

Setze $\Delta^{\text{alt}} \leftarrow \Delta$ und führe eine Relaxation

$$(u_h, \gamma_h) \leftarrow G_h(u_h, \gamma_h)$$

aus. Berechne dabei Δ analog zu Schritt b).

Falls $\Delta \leq \epsilon_k$ gilt, gehe nach e).

Ist dies nicht der Fall, gilt jedoch $\Delta < \eta \Delta^{\text{alt}}$, wiederhole Schritt d).

Ansonsten gehe im Fall $k > 0$ nach c), im Fall $k = 0$ wiederhole d).

e) Wechsel zum nächst feineren Gitter.

Setze $h \leftarrow h/2$, $k \leftarrow k+1$ und korrigiere

$$\begin{aligned} \underline{u}_h &\leftarrow \max\{0, \underline{u}_h + I_{2h}^h (\underline{u}_{2h} - \tilde{I}_h^{2h} \underline{u}_h)\} \\ \underline{\gamma}_h &\leftarrow \min\{1, \max\{0, \underline{\gamma}_h + I_{2h}^h (\underline{\gamma}_{2h} - \tilde{I}_h^{2h} \underline{\gamma}_h)\}\}. \end{aligned}$$

Falls $k < k_{\max}$ gilt, setze $\Delta \leftarrow 10^{30}$ und gehe nach d). Ansonsten fahre fort mit Schritt f).

f) A-posteriori-Glättung.

Führe zwei Relaxationen

$$(u_h, \gamma_h) \leftarrow G_h(u_h, \gamma_h)$$

aus.

Die automatische Steuerung von FASDAM wurde von Mehrgitter-Algorithmen zur Lösung einfacherer elliptischer Probleme übernommen. (Siehe "Cycle C" in [7], "PFAS" in [10]). Der Steuerparameter η orientiert sich am Glättungsfaktor der Relaxation, δ am Zweigitter-Konvergenzfaktor [7]. Um die Idee dieser Steuerung zu verdeutlichen, erläutern wir Schritt d) des obigen Algorithmus:

Nach einer Relaxation wird zunächst überprüft, ob die auf dieser Ebene angestrebte Genauigkeit ε_k erreicht ist. Falls ja, kann zum nächst feineren Gitter gewechselt werden. Im anderen Fall wird die "Konvergenzrate" der Relaxation geprüft. Ist sie noch befriedigend, so wird weiter relaxiert. Ist sie schlecht, so ist der Fehler bereits glatt, und die Iteration wird, soweit vorhanden, auf dem nächst gröberen Gitter fortgesetzt.

Ausdrücklich sei darauf hingewiesen, daß es unzumutbar ist, γ_h in irgendeiner Form zu glätten; so führt etwa eine Unterrelaxation von γ_h zu schlechterem Konvergenzverhalten.

5.4 Das Ende-Kriterium

Es bezeichne $(u_h, \gamma_h) \in M_h(u_h^0) \times L_h$ eine Näherung für die Lösung des Problems (1.2), die durch Iteration mit FASDAM gewonnen wurde. Ferner sei

$$\underline{d}_h := -A_h u_h - E_h \gamma_h,$$

$$\underline{d}_h^- := (d_h^{-1}, \dots, d_h^{-n}) \text{ mit}$$

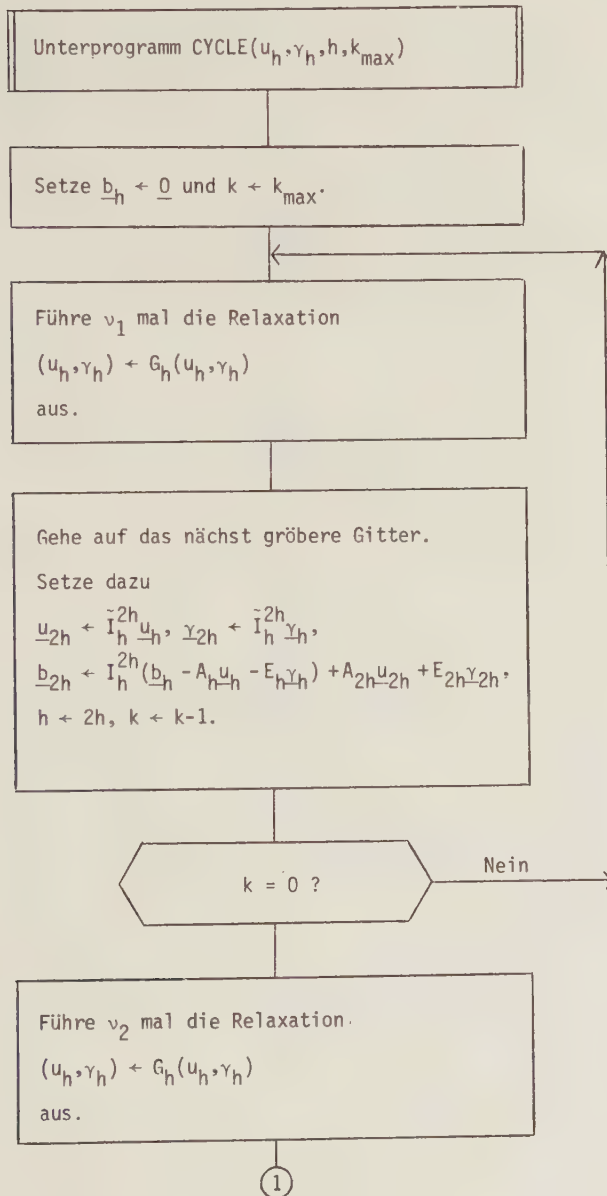
$$d_h^{-i} := \begin{cases} d_h^i & \text{falls } P_i \in D_h \text{ oder } [P_i \in P_h^0 \text{ und } \gamma_h^i < 1], \\ \min\{d_h^i, 0\} & \text{falls } P_i \in P_h^0 \text{ und } \gamma_h^i = 1, \\ 0 & \text{falls } P_i \in P_h^+. \end{cases}$$

Dann ist (u_h, γ_h) nach Folgerung 3.4 genau dann Lösung von (1.2), wenn $|d_h^-| = 0$ gilt. Allgemein ist die Größe $|d_h^-|$ ein anschauliches Maß für die Güte der Approximation (u_h, γ_h) . Da keine strengen Fehlerabschätzungen zur Verfügung stehen, wird der Algorithmus FASDAM beendet, wenn $h|d_h^-|$ kleiner als eine vorgeschriebene Toleranz ε ist.

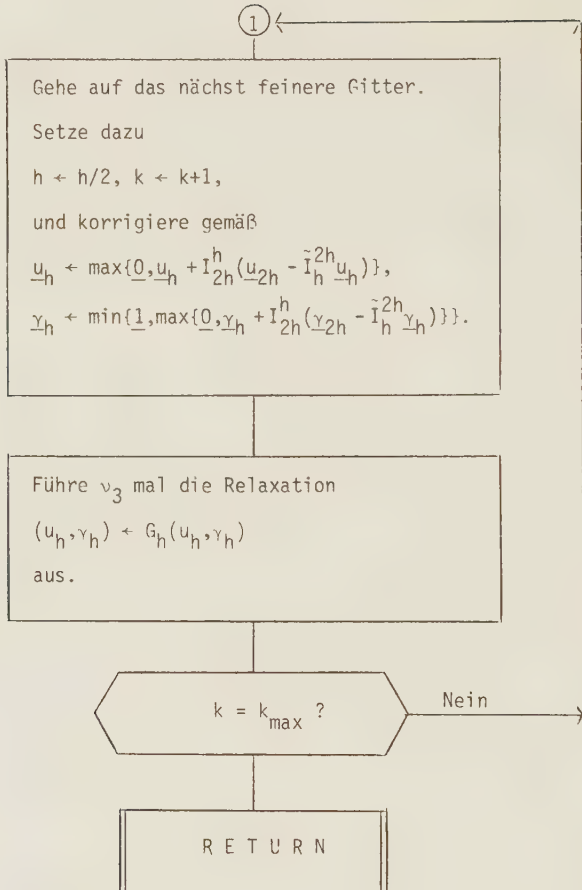
5.5 Der Full-Multigrid-Algorithmus

Auch beim Mehrgitter-Algorithmus mit voller Grob-Gitter-Korrektur erweist sich die Berechnung geeigneter Startwerte mit Hilfe der "Full Multigrid" Technik als vorteilhaft. D.h. die Startwerte für das Ω_{h_k} -Gitter werden auf dem $\Omega_{h_{k-1}}$ -Gitter berechnet usw.. Dabei wird auf jeder Stufe k solange mit Hilfe des entsprechenden Mehrgitterverfahrens iteriert, bis $h \cdot |d_h^-|$ kleiner als eine vorgeschriebene Toleranz ε_k ist.

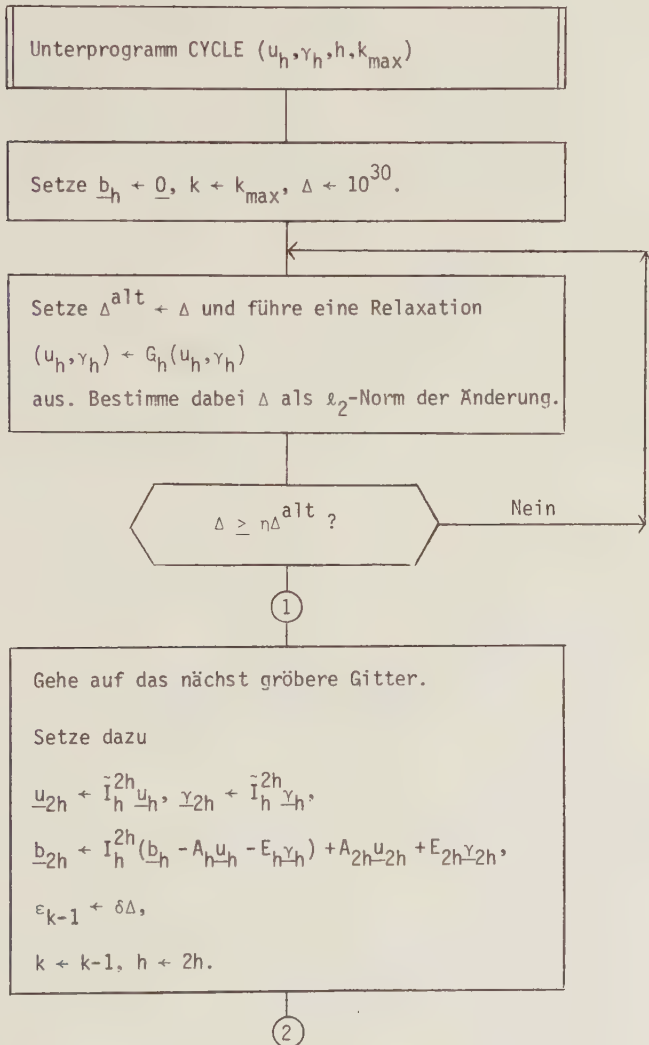
Flußdiagramm 4: Ein Iterationsschritt von FASDAM ohne automatische Steuerung (V-Zyklus)



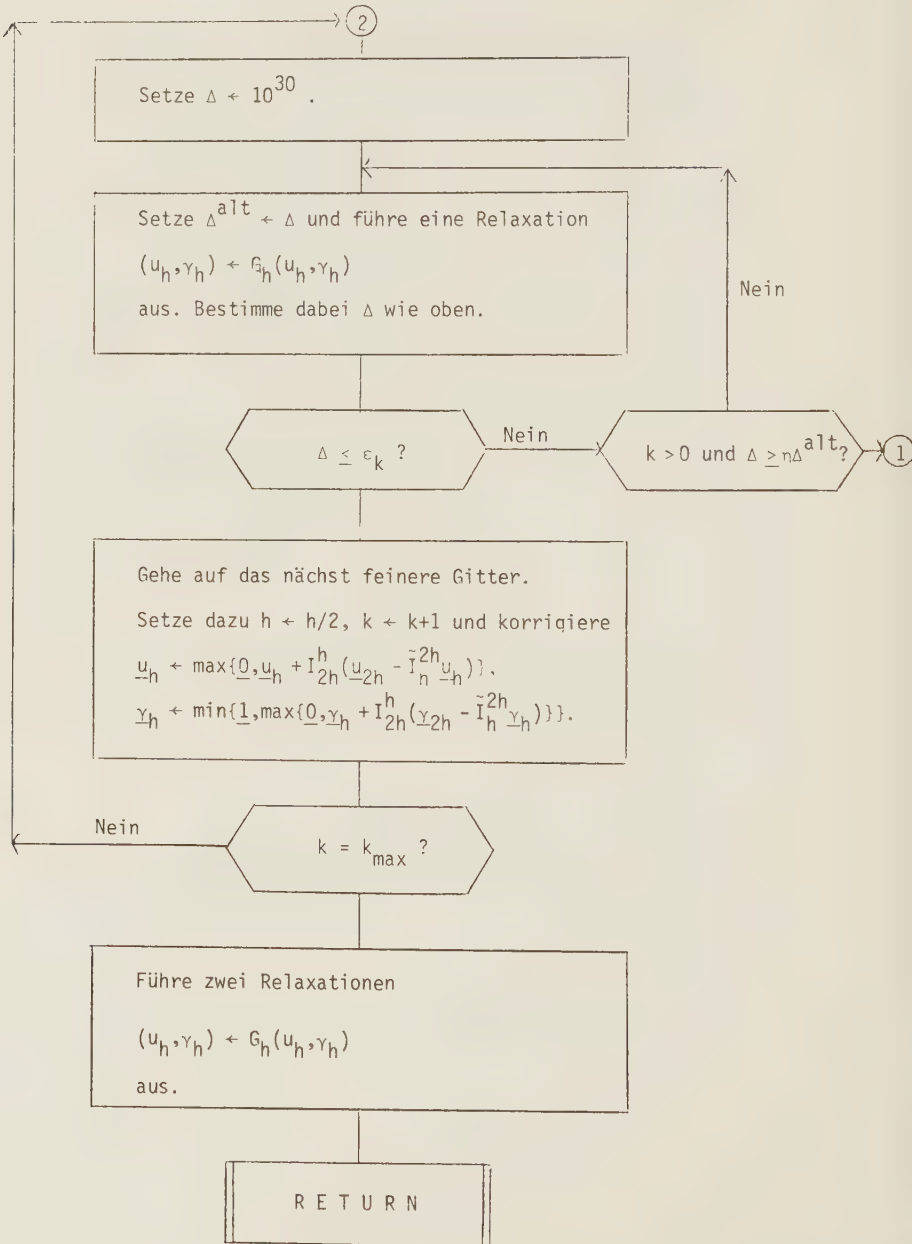
Fortsetzung Flußdiagramm 4



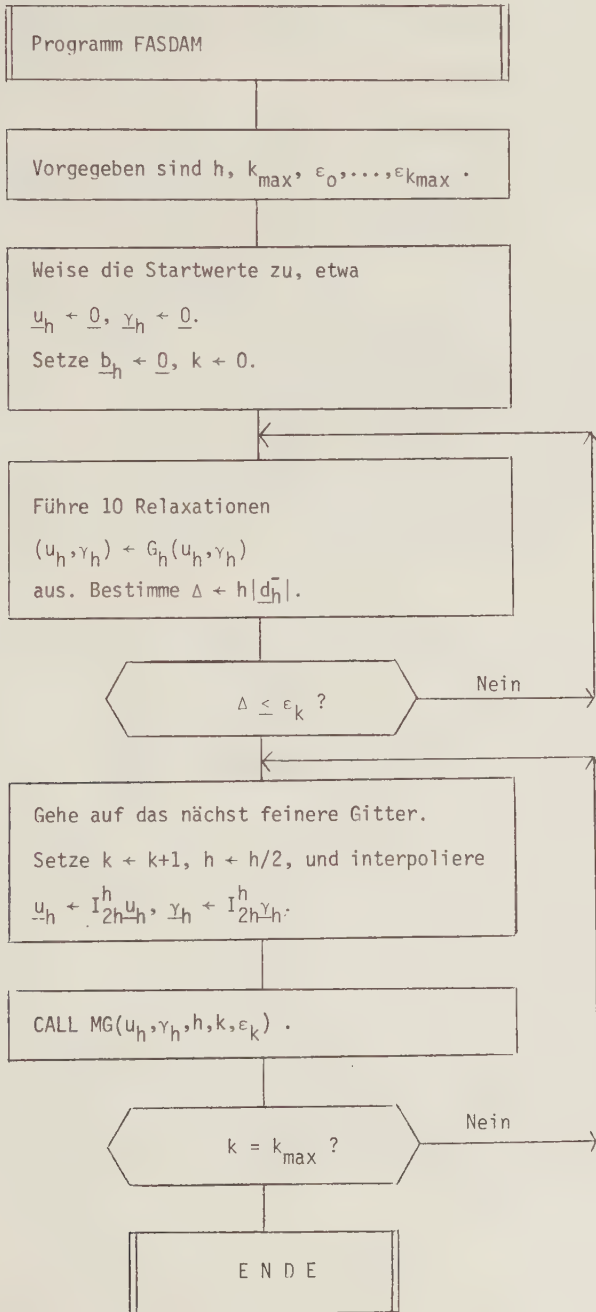
Flußdiagramm 5: Ein Zyklus von FASDAM mit automatischer Steuerung



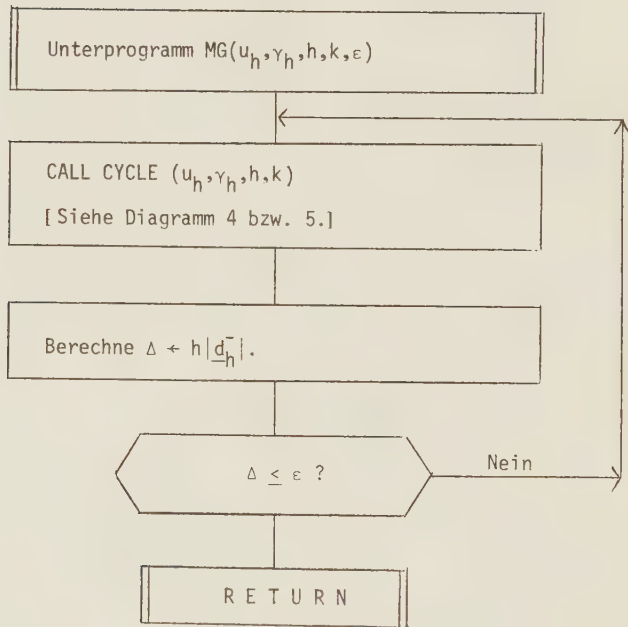
Fortsetzung Flußdiagramm 5



Flußdiagramm 6: Die Full Multigrid Version von FASDAM



Fortsetzung Flußdiagramm 6



6. Numerische Beispiele

6.1 Vergleich verschiedener Verfahren bei der Lösung eines Dammproblems mit quadratischem Damm

Das Beispiel 0 ist ein homogener isotroper Damm ($a_h = 1$) mit Querschnitt 1×1 zwischen Wasserreservoirien der Höhe 1 und 0.25. Der untere Rand des Dammes ist undurchlässig, der obere und der obere Teil des rechten Randes gehören zu r^0 .

Die Diskretisierung dieses Problems erfolgte mit Hilfe eines regelmäßigen Netzes gleicher Maschenbreite und -höhe. In diesem Fall entspricht der diskrete Differentialoperator dem Standard-5-Punkte-Stern. Da bei diesem Beispiel auch kein zusätzlicher Aufwand durch die Form des Gebietes entsteht, ist die Programmierung der Algorithmen besonders einfach. Hier ist also der Vergleich der Verfahren ohne zusätzlichen problembedingten Overhead möglich.

Zur Berechnung von Beispiel 0 wurden fünf Programme erstellt:

- a) RELDAM (0.1) berechnet eine Näherung für die Lösung des oben beschriebenen Problems mit Hilfe des Einzelschrittverfahrens (3.5).
- b) RELDAM0 (0.1) entspricht dem Incorporated SORP (3.9). Der optimale Overrelaxationsparameter ω wurde experimentell bestimmt. Unabhängig von der Schrittweite h erwies sich die Wahl $\omega=1.95$ als günstig.
- c) MGDAM (0.1) ist eine Full Multigrid Version des Algorithmus MGDAM, wie sie in Kapitel 4 vorgestellt wurde, ohne automatische Steuerung. (Siehe Flußdiagramm 1 und 3.) Hier erwies sich die Wahl $v_1=5$, $v_2=20$, $v_3=4$, $\epsilon_{k_{\max}}^{-1} = \frac{1}{16}$, $\epsilon_{k_{\max}}^{-2} = \frac{1}{8}$, ... als günstig.

- d) FASDAM (0.1) ist eine Variante des Mehrgitter-Algorithmus FASDAM ohne automatische Steuerung. Sie wurde in Flußdiagramm 4 beschrieben. Die Startwerte werden auch hier mit Hilfe der Full Multigrid Technik bestimmt. Wir haben $v_1 = v_3 = 2$, $v_2 = 10$, $\epsilon_k = 10^{-3}$ für $k < k_{\max}$ gewählt. Bei einer Verdoppelung von v_2 , der Anzahl der Relaxationsschritte auf dem größten Gitter, werden die unten aufgeführten Rechenzeiten geringfügig länger, wobei die Änderungen im Bereich von 10% liegen.
- e) FASDAM (0.2) entspricht dem vorstehenden Programm, die Anzahl der Relaxationen und Grob-Gitter-Korrekturen wird jedoch automatisch gesteuert. (Siehe Flußdiagramm 5 und 6.) Die Werte $\delta = 0.075$, $\eta = 0.5$ für die Steuerparameter wurden aus [10] übernommen. Das Programm reagiert auf Änderungen der Parameter nicht sehr sensibel.

Alle Rechnungen wurden auf der Cyber 175 des Rechenzentrums der Ruhr-Universität Bochum durchgeführt. Die Programme wurden in Fortran 77 geschrieben und mit der Optimierungsstufe 2 übersetzt.

Hier wie auch bei allen weiteren Beispielen haben wir zur Ermittlung des Fehlers bei den Versionen von FASDAM die "exakte" Lösung bis auf einen Fehler von 10^{-4} genau berechnet.

Das größte Gitter war in allen Fällen ein 5×5 Gitter.

Tabelle 1: Rechenzeiten und Fehler verschiedener Verfahren, Beispiel 0

Verfahren und Abbruch-Kriterium		9 × 9 Gitter	17 × 17 Gitter	33 × 33 Gitter	65 × 65 Gitter
RELDAM(0.1) $\ u_h^+ + y_h^+ - u_h^- - y_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0051	0.0092	0.0098	0.0099
	CPU-Zeit [s]	0.10	1.10	16.75	282.39
RELDAM(0.1) $\ u_h^+ + y_h^+ - u_h^- - y_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0075	0.0071	0.0076	0.0087
	CPU-Zeit [s]	0.06	0.34	2.84	27.12
MDGAM(0.1) $\ u_h^+ + y_h^+ - u_h^- - y_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0035	0.0091	0.0093	0.0095
	CPU-Zeit [s]	0.09	0.46	3.75	41.16
FASDAM(0.1) $h d_h^- \leq 10^{-6}$	Maximumnorm des Fehlers	0.0002	0.0002	0.0006	0.0030
	CPU-Zeit [s]	0.07	0.17	0.55	1.63
FASDAM(0.2) $h d_h^- \leq 10^{-6}$	Maximumnorm des Fehlers	0.0001	0.0004	0.0017	0.0030
	CPU-Zeit [s]	0.08	0.21	0.62	1.80

Bis auf den Fall der größten Triangulierung sind die Mehrgitter-Programme mit voller Grob-Gitter-Korrektur schneller als die anderen Algorithmen. Im Fall des 65 × 65 Gitters benötigen sie noch ca. 1% der Rechenzeit des Einzelschrittverfahrens und ca. 13% der Rechenzeit des Incorporated SORP. Das in Kapitel 4 entwickelte Verfahren MGDAM ist bis zu fast 50% langsamer als das Incorporated SORP. Dieses Phänomen ist durch den großen Rechenaufwand in Schritt c) von MGDAM (Berechnung einer Korrektur für u_h in den Punkten der gesättigten Zone) zu erklären. So muß bei jeder Durchführung dieses Schrittes die Menge $N_{2h}(y_h)$ der aktiven Punkte neu bestimmt werden. Bei einem Vergleich der beiden Verfahren sollte jedoch beachtet werden, daß MGDAM eine strenge Einschließung der Lösung und so Fehlerabschätzungen liefert, während das Incorporated SORP wie FASDAM ein heuristisches Verfahren darstellt.

Bei Verdoppelung der Schrittweite nimmt der Rechenaufwand bei FASDAM um den Faktor 3 zu. Dieses Verhalten ist typisch für Full-Multigrid-Verfahren dieser Größenordnung [8]. Bei RELDAM bzw. RELDAMO betragen die Faktoren 16 bzw. 8.

Um die Sensibilität der Konvergenzfaktoren von FASDAM gegenüber Änderungen in h zu untersuchen, haben wir ein schärferes Abbruch-Kriterium gewählt, so daß auf dem feinsten Gitter zwischen 9 und 13 Iterationen ausgeführt wurden, und die durchschnittlichen Dämpfungsfaktoren von $|\underline{d}_h^-|$ berechnet. Bei dem Mehrgitterverfahren mit fester Anzahl von Glättungen liegen die Dämpfungsfaktoren pro Zyklus in den ersten beiden Schritten typischerweise unter den durchschnittlichen Dämpfungsfaktoren, steigen dann rasch an und werden später langsam wieder kleiner, wobei nach den ersten Schritten Schwankungen von insgesamt ca. 10% auftreten. Bei FASDAM(0.2) schwanken die Dämpfungsfaktoren von Schritt zu Schritt um bis zu 50%.

Die automatische Steuerung führt bei den drei kleineren Problemen zu einem V-Zyklus, bei dem größten Problem entweder zu einem V-Zyklus oder zu einer Mischung aus V- und W-Zyklus, die einem umgekehrten F-Zyklus ähnelt. (Siehe dazu [9] und [27].) Da der Rechenaufwand pro Zyklus schwankt, wurden auch die durchschnittlichen Dämpfungsfaktoren pro Work Unit [WU] berechnet. Allgemein wird unter einem WU der Rechenaufwand eines Relaxationsschrittes auf dem feinsten Gitter verstanden. Jede Ausführung von G_{h_k} und jede Berechnung von $\underline{b}_{h_{k-1}}$ benötigt $4^{(k-k_{\max})}$ WU. Der Aufwand für Interpolation und Restriktion ist demgegenüber klein und wird wie üblich nicht berücksichtigt. Auch der bei der automatischen Steuerung zusätzlich anfallende Aufwand, die Berechnung von Kontrollgrößen usw., wird vernachlässigt. Das führt dazu, daß ein Work Unit ein recht grobes Maß für die verrichtete Arbeit darstellt. So beträgt bei der Berechnung von Beispiel 0 auf dem

65 × 65 Gitter mit Hilfe von FASDAM(0.2) das Verhältnis von Gesamtrechnenzeit zur Zahl der verrichteten WU 0.051, bei FASDAM(0.1) dagegen nur 0.046.

Tabelle 2: Vergleich der Dämpfungsfaktoren von FASDAM(0.1) und FASDAM(0.2)

Verfahren und Abbruch-Kriterium		9 × 9 Gitter	17 × 17 Gitter	33 × 33 Gitter	65 × 65 Gitter
FASDAM(0.1) $h \underline{d}_h^- \leq 10^{-10}$	ØDämpfungsfaktor von $ \underline{d}_h^- $ pro Zyklus	0.150	0.206	0.253	0.335
	ØDämpfungsfaktor von $ \underline{d}_h^- $ pro WU	0.777	0.795	0.815	0.849
FASDAM(0.2) $h \underline{d}_h^- \leq 10^{-10}$	ØDämpfungsfaktor von $ \underline{d}_h^- $ pro Zyklus	0.146	0.250	0.246	0.272
	ØDämpfungsfaktor von $ \underline{d}_h^- $ pro WU	0.763	0.816	0.831	0.846

Insgesamt ist festzuhalten, daß sich die durchschnittlichen Dämpfungsfaktoren pro WU bei einer Verachtfachung der Schrittweite nur um ca. 10% ändern.

Die Lösungen zu den Beispielen werden im Anhang dieser Arbeit anhand einiger Plots dargestellt. Dabei kennzeichnen Pfeile die Geschwindigkeit des sickern- den Wassers in der gesättigten Zone. Der Schwerpunkt jedes Dreiecks stimmt mit dem Mittelpunkt eines Pfeiles überein. Die Länge der Pfeile entspricht, logarithmisch skaliert, dem Betrag der Geschwindigkeit. Die ungesättigte Zone ist durch vertikale Schraffur angedeutet. Die Länge der Striche entspricht der Größe von γ_h im jeweiligen Dreieck. Der eingezeichnete freie Rand ist eine stetige Näherung für die Nulllinie von $u_h + \gamma_h$ und beschreibt, unserer Erfahrung nach, den in der Realität auftretenden freien Rand besser als etwa der Rand von $\{u_h > 0\}$. Die senkrecht zur Strömung verlaufenden Potentiallinien sind Linien gleicher Standrohrspiegelhöhe $u_h(\underline{x}) + \underline{x} \underline{e}$.

6.2 Weitere Beispiele

Wir haben die in den Kapiteln 3,4 und 5 vorgestellten Algorithmen bei der Berechnung einiger weiterer Beispiele verglichen. Die zu diesem Zweck erstellten Programme können zur Lösung beliebiger inhomogener isotroper Probleme herangezogen werden, vorausgesetzt Ω ist durch ein Netz der folgenden Art konform triangulierbar:

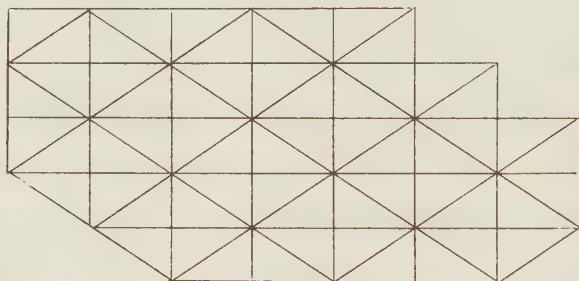


Fig. 5

Dabei können Maschenbreite und -höhe beliebig gewählt werden.

In diesem Abschnitt vergleichen wir vier Programme:

- f) `RELDAM(0.2)` stellt eine Erweiterung von `RELDAM(0.1)` für allgemeinere Probleme dar.
- g) `RELDAMO(0.2)` erweitert in gleicher Hinsicht das Programm `RELDAMO(0.1)`. Die optimalen Overrelaxationsparameter wurden für jedes Problem experimentell bestimmt. Sie sind i.a. unabhängig von der Schrittweite h und liegen nahe bei 2.

- h) MGDAM(0.2) ist eine Full Multigrid Version des in Kapitel 4 entworfenen Algorithmus MGDAM mit automatischer Steuerung. Die Wahl $\eta = 0.95$, $\delta = 0.25$, $\varepsilon_0 = 1.0$, $\varepsilon_1 = 0.5, \dots$ erwies sich als allgemein günstig. Lediglich bei den Beispielen 1 und 1a konnten wir analog zu Beispiel 0 die Werte $\varepsilon_{k_{\max}-1} = 0.25$, $\varepsilon_{k_{\max}-2} = 0.5, \dots$ wählen.
- i) FASDAM(0.3) stellt eine Erweiterung des Programms FASDAM(0.2) für allgemeinere Probleme dar. Bei Beispiel 3 wurde δ verdoppelt und damit der schlechten Konvergenzrate angepaßt.

Bei dem Beispiel 0 in 6.1 konnte man beobachten, daß das Mehrgitterverfahren mit automatischer Steuerung trotz vergleichbarer Konvergenzfaktoren aufgrund der zusätzlichen Berechnung von Kontrollgrößen langsamer war als die entsprechende Variante mit fester Anzahl von Grob-Gitter-Korrekturen und Relaxationen. Die optimale Anzahl von Relaxationen schwankt jedoch von Beispiel zu Beispiel erheblich, und das Programm reagiert auf Änderungen zum Teil sehr sensibel. Daher wird im folgenden bei FASDAM wie bei MGDAM nur die Variante mit automatischer Steuerung betrachtet.

Die in diesem Kapitel behandelten Beispiele wurden in ähnlicher Form bereits von H.W. Alt und L.D. Marini & P. Pietra zum Test ihrer Verfahren herangezogen.

Beispiel 1

Das Beispiel 1 ist ein homogener isotroper Damm von trapezförmigem Querschnitt zwischen zwei Wasserreservoirien der Höhe 1 und 0.25. Die Höhe des Trapezes beträgt 1, die Länge der Mittellinie 2. Der untere Rand des Damms ist undurchlässig. Der verbleibende Teil des Randes gehört zu Γ^0 .

Das gröbste Gitter besteht in jedem Fall aus 45 Knoten. Die Maschenbreite ist gleich der Maschenhöhe.

Tabelle 3: Rechenzeiten und Fehler verschiedener Verfahren, Beispiel 1.

Verfahren und Abbruch-Kriterium		153 Knoten	561 Knoten	2145 Knoten
RELDAM(0.2) $\ u_h^+ + \gamma_h^+ - u_h^- - \gamma_h^-\ _\infty \leq 10^{-2}$ *	Maximumnorm des Fehlers	0.0098	0.0098	0.0198
	CPU-Zeit [s]	2.58	38.29	543.87
RELDAM0(0.2) $\ u_h^+ + \gamma_h^+ - u_h^- - \gamma_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0086	0.0093	0.0088
	CPU-Zeit [s]	0.77	5.59	47.03
MGDAM(0.2) $\ u_h^+ + \gamma_h^+ - u_h^- - \gamma_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0086	0.0093	0.0099
	CPU-Zeit [s]	1.12	8.84	87.96
FASDAM(0.3) $h d_h^- \leq 10^{-3}$	Maximumnorm des Fehlers	0.0096	0.0034	0.0035
	CPU-Zeit [s]	0.18	0.52	1.51

* Um Rechenzeit zu sparen, wurde beim größten Problem das Abbruch-Kriterium

$$\|u_h^+ + \gamma_h^+ - u_h^- - \gamma_h^-\|_\infty \leq 0.02 \text{ gewählt.}$$

Tabelle 4: Dämpfungsfaktoren von FASDAM(0.3), Beispiel 1.

Verfahren und Abbruch-Kriterium		153 Knoten	561 Knoten	2145 Knoten
FASDAM (0.3) $h d_h^- \leq 10^{-10}$	ϕ Dämpfungsfaktor von $ d_h^- $ pro Zyklus	0.207	0.266	0.256
	ϕ Dämpfungsfaktor von $ d_h^- $ pro WU	0.900	0.876	0.863

Im wesentlichen ergibt sich dasselbe Bild wie bei Beispiel 0. FASDAM erscheint bei diesem komplizierteren Problem gegenüber den anderen Verfahren

noch günstiger. Die Dämpfungsfaktoren pro WU sind bei feiner werdender Diskretisierung und dabei steigender Zahl verwendeter Gitter (unabhängig von h) von 1 weg beschränkt.

Als nächstes haben wir zwei inhomogene Varianten von Beispiel 1 untersucht. Bei Beispiel 1a beträgt die Permeabilität in der rechten Dammhälfte 5, in der linken Hälfte 1. Bei Beispiel 1b ist die Situation umgekehrt.

Tabelle 5: Rechenzeiten und Fehler verschiedener Verfahren bei den Beispielen 1, 1a, 1b.

Verfahren und Abbruch-Kriterium		Beispiel 1 561 Knoten	Beispiel 1a 561 Knoten	Beispiel 1b 561 Knoten
RELDAM(0.2) $\ u_h^+ + y_h^+ - u_h^- - y_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0098	0.0099	0.0099
	CPU-Zeit [s]	38.29	55.78	27.96
RELDAMO(0.2) $\ u_h^+ + y_h^+ - u_h^- - y_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0093	0.0096	0.0063
	CPU-Zeit [s]	5.59	10.83	3.88
MGDAM(0.2) $\ u_h^+ + y_h^+ - u_h^- - y_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0093	0.0099	0.0079
	CPU-Zeit [s]	8.84	13.67	7.04
FASDAM(0.3) $h d_h^- \leq 10^{-3}$	Maximumnorm des Fehlers	0.0034	0.0054	0.0020
	CPU-Zeit [s]	0.52	0.58	0.50

Tabelle 6: Vergleich der Dämpfungsfaktoren von FASDAM(0.3) bei den Beispielen 1, 1a, 1b.

Verfahren und Abbruch-Kriterium		Beispiel 1 561 Knoten	Beispiel 1a 561 Knoten	Beispiel 1b 561 Knoten
FASDAM(0.3) $h d_h^- \leq 10^{-10}$	ϕ Dämpfungsfaktor von $ d_h^- $ pro Zyklus	0.266	0.271	0.275
	ϕ Dämpfungsfaktor von $ d_h^- $ pro WU	0.876	0.894	0.859

Hier zeigt sich, daß das Mehrgitterverfahren mit voller Grob-Gitter-Korrektur auf Änderungen der Bodendurchlässigkeit und damit verbundene Verschiebungen des freien Randes bedeutend unempfindlicher reagiert als die beiden Relaxationsverfahren oder auch MGDAM.

Beispiel 2

Bei diesem Beispiel beträgt die maximale Höhe des porösen Mediums 1, die maximale Breite 1.75. Die Gestalt von Ω entnehme man dem Plot im Anhang. Der "Damm" sei homogen und isotrop. Wählt man die linke untere Ecke von Ω als Koordinatenursprung, so liegt links von Ω auf der Höhe $0.75 \leq x_2 \leq 1$ ein Wasserreservoir. Der obere Teil von Γ gehört wie der untere Rand im linken Teil zu Γ^0 . Alle anderen Randstücke seien undurchlässig. In dieser Situation ist die Lösung des diskreten Dammproblems (1.2) wie die Lösung des kontinuierlichen Problems (1.1) nicht eindeutig bestimmt. Addiert man zu der geplotteten (minimalen) Lösung (u_h, γ_h) ein Element w_h^C der Menge

$$(6.1) \quad \{w_h^C : \Omega \rightarrow \mathbb{R} \times \mathbb{R}, w_h^C(\cdot) = \chi(\{0.5 < x_1 < 1.75, 0 < x_2 < c\})(c - \underline{e}, 1), 0 \leq c \leq 0.5\},$$

so erhält man eine weitere Lösung des Problems (1.2). Im Fall $c \neq 0$ treten dann Zusammenhangskomponenten von $\{u_h > 0\}$ auf, die nicht in Kontakt zu Γ^+ stehen. Diese Zusammenhangskomponenten werden als Grundwasserreservoirs ("ground water reservoirs", [4]) oder als Becken ("pools", [12]) bezeichnet. Es ist bekannt, daß im homogenen isotropen Fall die Lösung des kontinuierlichen Dammproblems bis auf diese Becken eindeutig bestimmt ist (Theorem 9.3 in [4], Theoreme 4.4 und 5.2 in [12]).

Löst man das oben beschriebene Problem mit Hilfe von konvergenten Folgen von Ober- und Unterlösungen, so unterscheiden sich die beiden Grenzwerte genau um das Element $w_h^{0.5}$ aus der Menge (6.1). Um trotzdem strenge Fehler-

abschätzungen zu erhalten, haben wir, sobald zwei Zusammenhangskomponenten von $\{u_h^+ > 0\}$ auftraten, im rechten Teil des "Damms" $(u_h^+, y_h^+) \leftarrow (0, 0)$ gesetzt. Man beachte, daß bei diesem Vorgehen die Eigenschaft, Oberlösung zu sein, erhalten bleibt.

Als Ω_h wählten wir ein regelmäßiges Netz gleicher Maschenbreite und -höhe. Das größte verwendete Gitter enthält 33 Knoten. Da das Problem auf diesem Gitter eine eindeutige diskrete Lösung besitzt, beginnen die Full Multigrid Programme erst auf Ω_{h_1} .

Tabelle 7: Rechenzeiten und Fehler verschiedener Verfahren, Beispiel 2.

Verfahren und Abbruch-Kriterium		108 Knoten	387 Knoten	1425 Knoten
RELDAM(0.2) $\ u_h^+ + y_h^+ - u_h^- - y_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0054	0.0077	0.00001
	CPU-Zeit [s]	2.37	25.00	412.95
RELDAM(0.2) $\ u_h^+ + y_h^+ - u_h^- - y_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0055	0.0067	0.0084
	CPU-Zeit [s]	0.94	6.11	43.92
MGDAM(0.2) $\ u_h^+ + y_h^+ - u_h^- - y_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0014	0.0069	0.0095
	CPU-Zeit [s]	2.40	6.09	42.05
FASDAM(0.3) $h d_h^- \leq 10^{-3}$	Maximumnorm des Fehlers	0.0039	0.0016	0.0008
	CPU-Zeit [s]	0.08	0.28	1.03

Tabelle 8: Dämpfungsfaktoren von FASDAM(0.3), Beispiel 2.

Verfahren und Abbruch-Kriterium		108 Knoten	387 Knoten	1425 Knoten
FASDAM(0.3) $h d_h^- \leq 10^{-10}$	ϕ Dämpfungsfaktor von $ d_h^- $ pro Zyklus	0.091	0.112	0.125
	ϕ Dämpfungsfaktor von $ d_h^- $ pro WU	0.670	0.714	0.689

Die kurzen Rechenzeiten und kleinen Konvergenzfaktoren sind auf die Tatsache zurückzuführen, daß es sich hier bei Nichtachtung des Beckens um sehr kleine Probleme handelt.

Beispiel 3

Durch eine geringfügige Veränderung der Randbedingung am linken Rand von Beispiel 2 - das Reservoir beginnt jetzt bereits auf der Höhe 0.5 - erhält man ein Problem mit eindeutiger Lösung. Bei gleicher Diskretisierung gelangten wir zu folgenden Ergebnissen:

Tabelle 9: Rechenzeiten und Fehler verschiedener Verfahren, Beispiel 3.

Verfahren und Abbruch-Kriterium		108 Knoten	387 Knoten
RELDAM(0.2) $\ u_h^+ + \gamma_h^+ - u_h^- - \gamma_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0100	0.0099
	CPU-Zeit [s]	7.12	97.33
RELDAM(0.2) $\ u_h^+ + \gamma_h^+ - u_h^- - \gamma_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0089	0.0095
	CPU-Zeit [s]	2.82	19.87
MGDAM(0.2) $\ u_h^+ + \gamma_h^+ - u_h^- - \gamma_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0100	0.0100
	CPU-Zeit [s]	5.91	41.77
FASDAM(0.3) $h d_h^- \leq 10^{-4}$	Maximumnorm des Fehlers	0.0036	0.0010
	CPU-Zeit [s]	0.41	0.86

Tabelle 10: Dämpfungsfaktoren von FASDAM(0.3), Beispiel 3.

Verfahren und Abbruch-Kriterium		108 Knoten	387 Knoten
FASDAM(0.3) $h d_h^- \leq 10^{-10}$	ϕ Dämpfungsfaktor von $ d_h^- $ pro Zyklus	0.600	0.530
	ϕ Dämpfungsfaktor von $ d_h^- $ pro WU	0.958	0.940

Die Lösung ist im linken Teil von Ω durch große, im rechten Teil durch extrem kleine Sickergeschwindigkeiten charakterisiert. Diese Situation führt bei allen Verfahren zu langsamer Konvergenz. Auch hier reagiert das Mehrgitter-Programm mit voller Grob-Gitter-Korrektur weniger sensibel als die anderen Verfahren.

Beispiel 4

Bei diesem Problem sickert Wasser aus einem links oberhalb von Ω gelegenen Reservoir durch ein stark inhomogenes Medium; in der oberen Hälfte beträgt die Permeabilität 1, in der unteren Hälfte 8. Der linke Rand von Ω ist undurchlässig und kann als Symmetrieachse interpretiert werden. Der untere Rand ist ebenfalls undurchlässig. Der rechte Rand und die rechte Hälfte des oberen Randes gehören zu Γ^0 . Theorem 4.4 in [2] besagt, daß sich $\Omega \cap \partial\{u > 0\}$ unter der Voraussetzung $\nabla(ae) \geq 0$ als Graph einer Funktion schreiben läßt. Hier ist diese Bedingung nicht erfüllt, und es treten zwei "freie Ränder" auf. Zwischen ihnen liegt eine breite ungesättigte Zone.

Die größte verwendete Triangulierung enthält 33 Knoten. Das Verhältnis von Maschenbreite zu Maschenhöhe ist 4 : 3.

Tabelle 11: Rechenzeiten und Fehler verschiedener Verfahren, Beispiel 4.

Verfahren und Abbruch-Kriterium		110 Knoten	399 Knoten	1517 Knoten
RELDAM(0.2) $\ u_h^+ + y_h^+ - u_h^- - y_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0100	0.0100	0.0100
	CPU-Zeit [s]	2.49	33.55	510.59
RELDAM0(0.2) $\ u_h^+ + y_h^+ - u_h^- - y_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	0.0098	0.0093	0.0096
	CPU-Zeit [s]	1.17	8.64	69.93
MGDAM(0.2) $\ u_h^+ + y_h^+ - u_h^- - y_h^-\ _\infty \leq 10^{-2}$	Maximumnorm des Fehlers	/		0.0100
	CPU-Zeit [s]			63.59
FASDAM(0.3) $h d_h^- \leq 10^{-3}$	Maximumnorm des Fehlers	0.0012	0.0023	0.0021
	CPU-Zeit [s]	0.16	0.35	1.00

Tabelle 12: Dämpfungsfaktoren von FASDAM(0.3), Beispiel 4.

Verfahren und Abbruch-Kriterium		110 Knoten	399 Knoten	1517 Knoten
FASDAM(0.3) $h d_h^- \leq 10^{-10}$	ϕ Dämpfungsfaktor von $ d_h^- $ pro Zyklus	0.304	0.229	0.289
	ϕ Dämpfungsfaktor von $ d_h^- $ pro WU	0.894	0.883	0.870

Während FASDAM sich ganz ähnlich wie im Beispiel 0 oder im Beispiel 1 verhält, erweist sich hier das Verfahren MGDAM als ungeeignet. Die Anwendung von MGDAM ist nur bei Problemen mit kleinem ungesättigten Bereich sinnvoll.

6.3 Berechnung der Grundwasserströmung durch einen Deich am linken Niederrhein

Die bisher vorgestellten Beispiele sind eher theoretischer Natur. Im Rahmen der Zusammenarbeit mit den ingenieurwissenschaftlichen Abteilungen der RUB wurden die Programme auch zur Berechnung eines realen Damms herangezogen. Eine Arbeitsgruppe von Bauingenieuren und Hydrogeologen untersucht den Einfluß von Bergsenkungen auf die Grundwassersituation am linken Niederrhein. Dabei trat die Frage auf, wieviel Wasser in einem Brunnen am Fuß des Rheindeichs gefördert werden muß, damit am Deich keine Hangquelle auftritt. Wir betrachten folgenden vertikalen Schnitt durch einen Grundwasserleiter, der bis auf geringfügige Vereinfachungen dem realen Querschnitt entspricht.

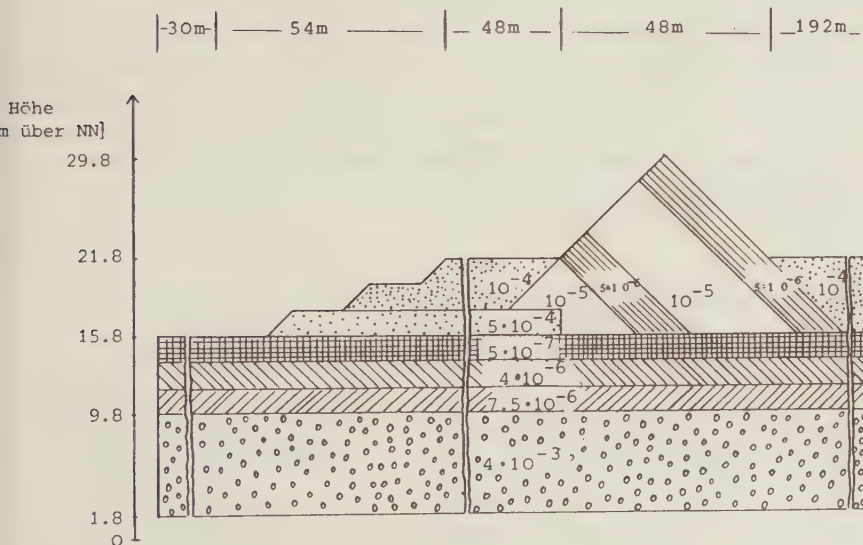


Fig. 6: Skizze des gerechneten Vertikalschnitts mit Werten für die Bodendurchlässigkeit [$\frac{m}{s}$]

An der rechten Dammseite steht das Wasser bis auf Höhe der Dammspitze. Die Unterkante des Querschnitts wird als undurchlässig betrachtet. Am linken Rand befindet sich ein Brunnen, mit dessen Hilfe der Dammfuß trocken gehalten werden soll. Zu fünf unterschiedlichen Potentialen in dem Brunnen wurden als Reaktionsmengen die entsprechenden Förderleistungen bestimmt. Anhand der geplotteten Darstellung der Grundwassersituation konnte entschieden werden, ob eine Hangquelle auftritt oder nicht.

Als gröbste Triangulierung verwendeten wir ein Netz mit 676 Knoten und einem Verhältnis von Maschenbreite zu Maschenhöhe von 3 : 1. Auf diesem Gitter berechneten wir die numerische Lösung mit Hilfe des Programms RELDAMO(0.2). Dabei wurde als Overrelaxationsparameter 1.8 gewählt. In den zwei interessanten Situationen verfeinerten wir das Netz auf 2560 Knoten und arbeiteten mit dem Programm FASDAM(0.3). Dabei entsprach die Wahl der Steuerparameter η und δ der unter 6.1. Außerdem wurde $\epsilon_0 = 10^{-6}$ und $\epsilon_1 = 10^{-7}$ gesetzt. Die Verwendung von mehr als zwei Gittern war aufgrund des sehr inhomogenen Bodens mit Schichten geringer Stärke und aufgrund der komplizierten Gestalt von Ω nicht möglich. (Siehe dazu die ergänzenden Bemerkungen in Kapitel 7.)

Tabelle 13: Pumpleistungen bei Beispiel 5.

(Die Werte in Klammern beziehen sich auf die feinere Diskretisierung mit 2560 Knoten.)

Potential im Brunnen [m über NN]	Pumpleistung [$\ell/s/m$]	Dammfuß	Flurabstand im Dammfuß [m]	CPU-Zeit [s]
5.80	2.10	trocken	ca. 5.6	248.22
7.80	1.98	trocken	ca. 4.8	237.77
9.80	1.83(1.81)	trocken(trocken)	ca. 3.2(3.6)	224.24(242.4)
11.80	1.65(1.63)	feucht (trocken)	≤ 0 . (0.8)	204.93(248.2)
13.80	1.47	feucht	≤ 0 .	185.41

Die Dämpfungsfaktoren von $|\underline{d}_h^-|$ pro Zyklus des Zweigitterverfahrens betragen 0.975 bzw. 0.971. Diese schlechten Werte sind auf das komplizierte Problem sowie auf die Verwendung von nur zwei Gittern zurückzuführen. So besitzt bei der Berechnung von Beispiel 0 auf einem 65×65 Gitter FASDAM(0.2) als Zweigitterverfahren auch nur den Dämpfungsfaktor 0.594 pro Zyklus (gegenüber 0.272 bei der Verwendung von 4 Gittern).

7. Ergänzende Bemerkungen

In der vorliegenden Arbeit haben wir uns auf die Behandlung eines recht einfachen Spezialfalls des in [3] vorgestellten Problems beschränkt. Mit der von H.W. Alt entwickelten schwachen Formulierung können etwa auch nichttriviale Neumann-Randbedingungen und Kapillareffekte erfaßt werden. Diese Einschränkungen sind jedoch technischer und nicht grundsätzlicher Natur. Bei entsprechender Triangulierung arbeitet der Kalkül aus [3] auch bei schwach anisotropen Problemen. Die zur Behandlung von Anisotropie bei Mehrgitterverfahren notwendigen Modifikationen ("line-relaxation", "semi-coarsening") wurden für lineare elliptische Probleme bereits ausführlich studiert. (Siehe etwa [9], [18], [21], [25].) Eine stabile Diskretisierung stark anisotroper Dämme ist nach Wissen des Verfassers noch unbekannt.

Alle unsere Programme gehen von einer regelmäßigen Triangulierung des Gebietes Ω aus. Verallgemeinerungen für den Fall eines unregelmäßigen Ω_h -Gitters sind mit Hilfe der Up-Wind Technik von P. Pietra möglich [26].

Grundsätzlich besteht auch die Möglichkeit, die Dreiecke des größten Gitters größer als die Bodenschichten zu wählen. Diese Nicht-Konformität sollte bei der Konstruktion des Mehrgitterverfahrens berücksichtigt werden. Die Interpolation etwa sollte im Sinne der "continuity interpolation" ([1], [22]) unter dem Gesichtspunkt der Stetigkeit von $v = \sum_i a_h [\nabla (I_{2h}^h u_{2h})^i \phi_h^i + (I_{2h}^h \gamma_{2h})^i x_h^i]$ konstruiert werden oder mit Hilfe des Relaxationsoperators R_h erfolgen. Entsprechend sind dann die Restriktion und die Matrizen A_{2h} und E_{2h} zu bestimmen.

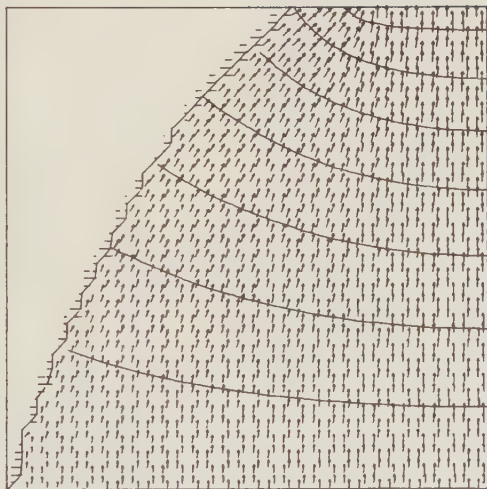
Stationäre zweidimensionale Dammprobleme können in der Regel mittels der

üblichen Relaxationsverfahren noch mit erträglichem Aufwand gelöst werden. Die Grenze der Aussagefähigkeit dieser Modelle wird in der Praxis nur allzu oft deutlich. Bei der Behandlung dreidimensionaler und/oder instationärer Probleme kann auf den Einsatz schneller Lösungsverfahren kaum verzichtet werden. Diese Probleme besitzen eine ähnliche Struktur wie das Problem (1.2), und ihre Behandlung mit Mehrgitterverfahren wie FASDAM sollte möglich sein.

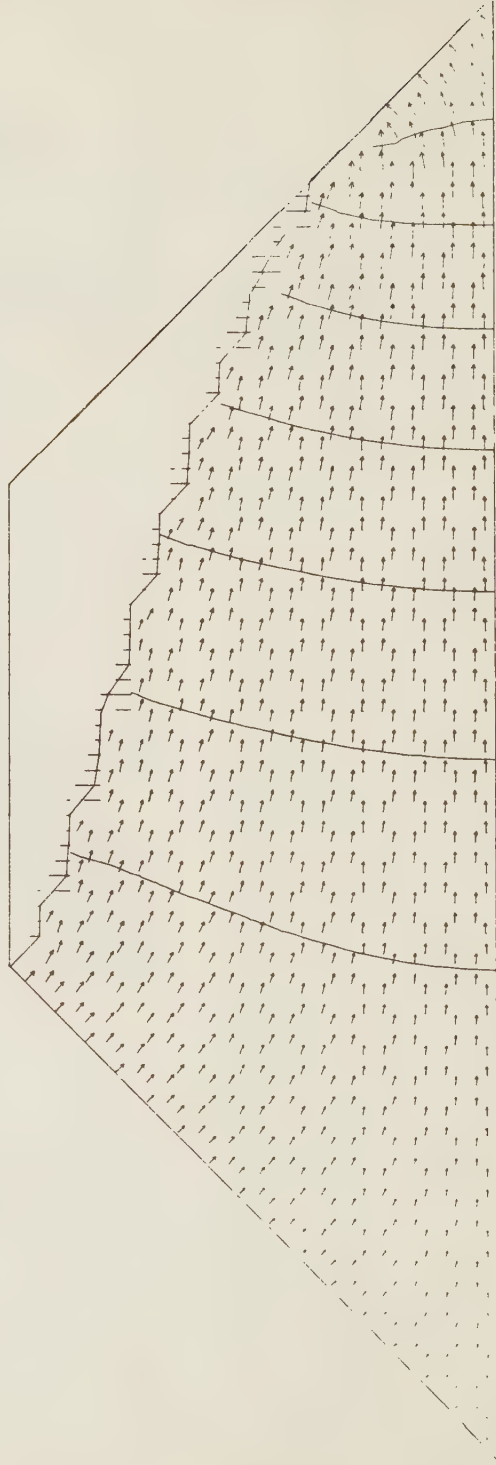
Anhang

Geplottete Darstellungen der Grundwasserströmung in den Beispielen 0 bis 4 aus den Abschnitten 6.1 und 6.2 sowie zweier Situationen in dem Damm an linken Niederrhein aus Abschnitt 6.3 (Beispiel 5).

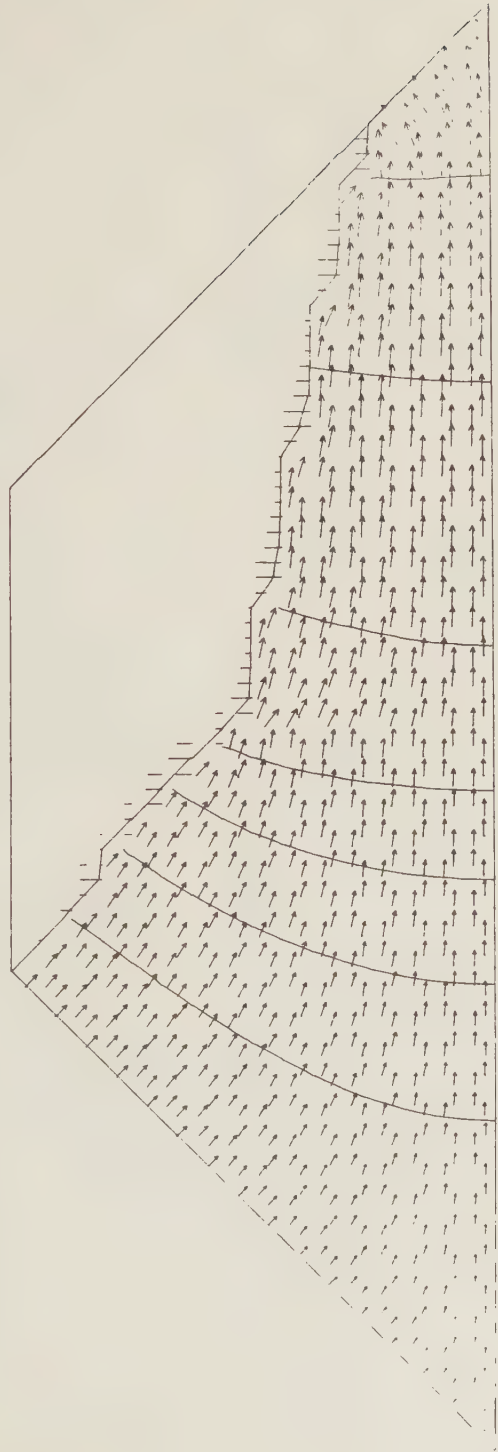
BEISPIELO 1089 KNOTEN



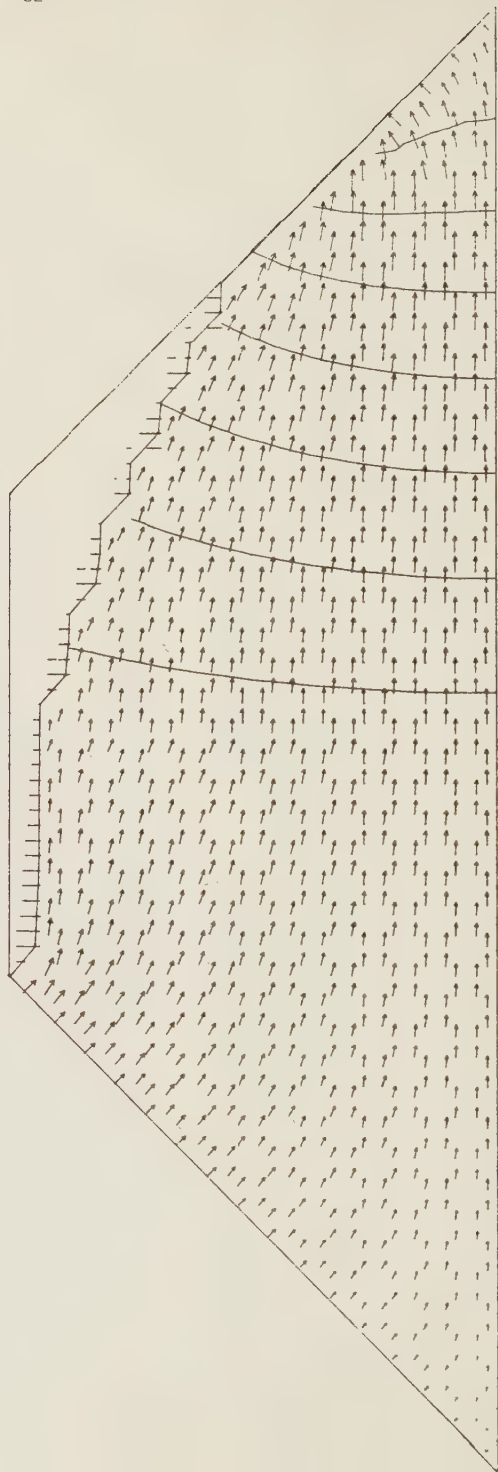
BEISPIEL 1 561 KNOTEN



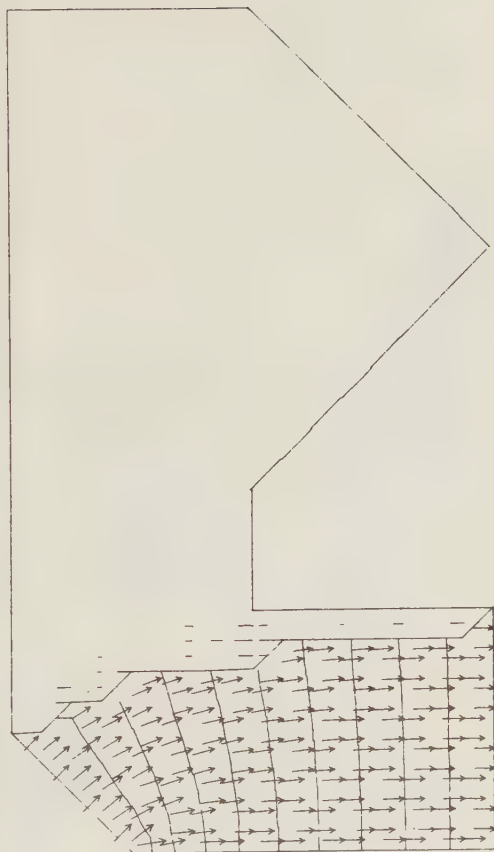
BEISPIEL 1A 561 KNOTEN



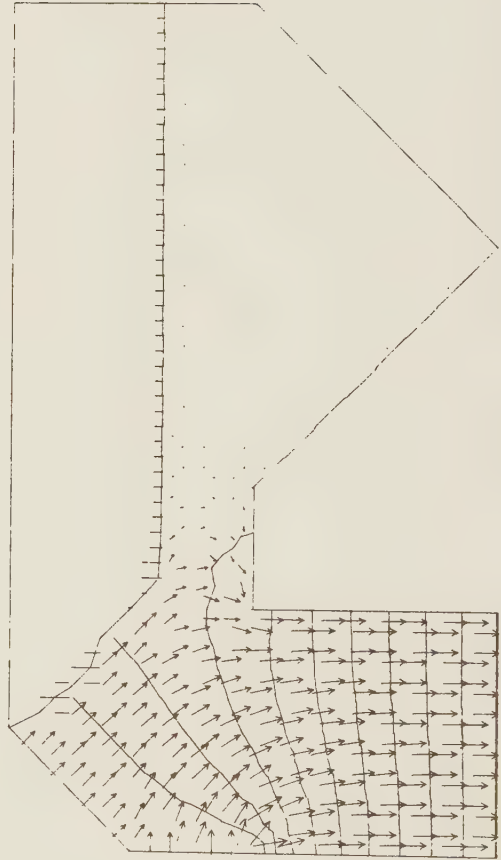
BEISPIEL 1B 561 KNOTEN



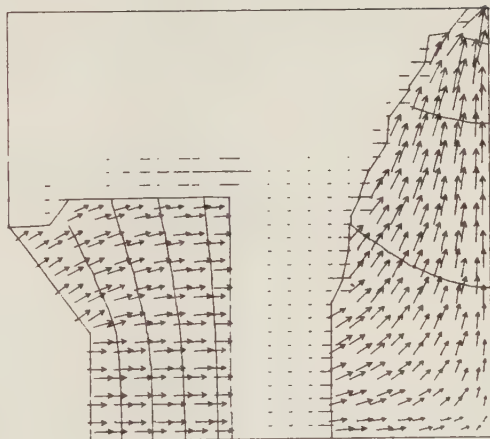
BEISPIEL 2 387 KNOTEN

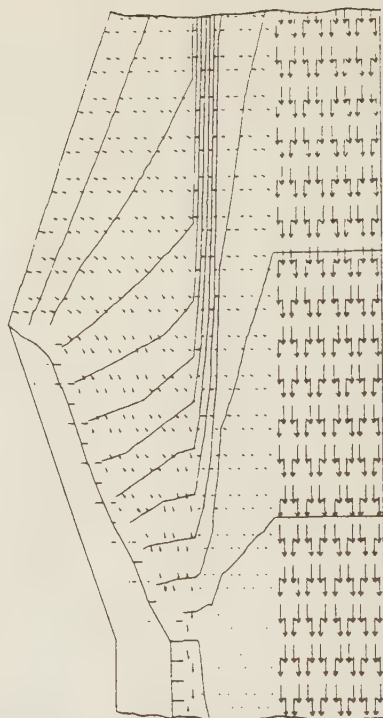
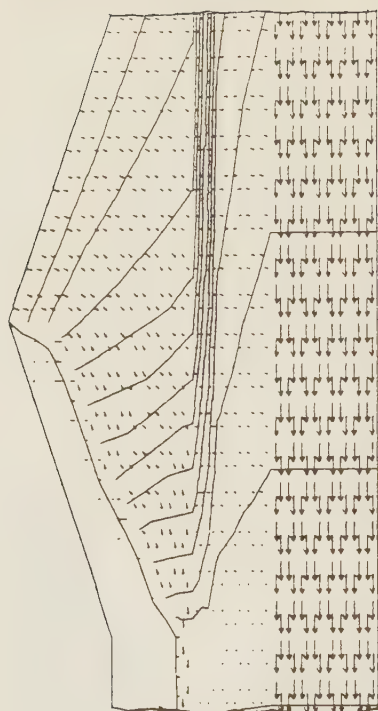


BEISPIEL 3 387 KNOTEN



BEISPIEL 4 399 KNOTEN





Literatur

- [1] R.E. Alcouffe, A. Brandt, J.E. Dendy Jr., J.W. Painter, The multi-grid method for the diffusion equation with strongly discontinuous coefficients, *SIAM J. Sci. Stat. Comput.* 2 (1981), 430-454.
- [2] H.W. Alt, Strömungen durch inhomogene poröse Medien mit freiem Rand, *J. Reine Angew. Math.* 305 (1979), 89-115.
- [3] H.W. Alt, Numerical solution of steady-state porous flow free boundary problems, *Numer. Math.* 36 (1980), 73-98.
- [4] H.W. Alt, G. Gilardi, The behaviour of the free boundary for the dam problem, *Ann. Sc. Norm. Sup. Pisa IX(4)* (1982), 571-626.
- [5] C. Baiocchi, V. Comincioli, L. Guerri, G. Volpi, Free boundary problems in the theory of fluid flow through porous media: a numerical approach, *Calcolo* 10 (1973), 1-86.
- [6] C. Baiocchi, V. Cominciolo, E. Magenes, G.A. Pozzi, Free boundary problems in the theory of fluid flow through porous media: existence and uniqueness theorems, *Ann. Mat. Pura Appl. (4)* 97 (1973), 1-82.
- [7] A. Brandt, Multi-level adaptive solutions to boundary value problems, *Math. Comp.* 31 (1977), 333-390.
- [8] A. Brandt, Multi-level adaptive finite-element methods. I. Variational problems. *Special Topics of Applied Mathematics*, J. Frehse, D. Pallaschke, U. Trottenberg (eds.), North-Holland 1980, 91-128.
- [9] A. Brandt, Guide to multigrid development, 220-312 in [21].
- [10] A. Brandt, C.W. Cryer, Multigrid algorithms for the solution of linear complementarity problems arising from free boundary problems, *SIAM J. Sci. Stat. Comput.* 4 (1983), 655-684.
- [11] H. Brezis, D. Kinderlehrer, G. Stampacchia, Sur une nouvelle formulation du problème de l'écoulement à travers une digue, *C.R. Acad. Sci. Paris* 287 (1978), 711-714.

- [12] J. Carrillo-Menendez, M. Chipot, On the dam problem, J. differential Equations 45 (1982), 234-271.
- [13] P.G. Ciarlet, The finite element method for elliptic problems, North-Holland 1981.
- [14] P.G. Ciarlet, P.A. Raviart, Maximum principle and uniform convergence for the finite element method, Comput. Methods Appl. Mech. Engrg. 2(1973), 17-31.
- [15] C.W. Cryer, The solution of a quadratic programming problem using systematic overrelaxation, SIAM J. Control 9 (1971), 385-392.
- [16] C.W. Cryer, A survey of steady-state porous flow free boundary problems, M.R.C. Univ. Wisconsin, 1657 (1976).
- [17] R.S. Falk, Error estimates for the approximation of a class of variational inequalities, Math. Comp. 28 (1974), 963-971.
- [18] H. Federer, Geometric measure theory, Springer 1969.
- [19] W. Hackbusch, Multi-grid convergence theory, 177-219 in [21].
- [20] W. Hackbusch, H.D. Mittelmann, On multigrid methods for variational inequalities, Numer. Math. 42 (1983), 65-76.
- [21] W. Hackbusch, U. Trottenberg (eds.), Multigrid methods, Springer 1982.
- [22] R. Kettler, Analysis and comparison of relaxation schemes in robust multigrid and preconditioned conjugate gradient methods, 502-534 in [21].
- [23] J. Mandel, A fast iterative method for large, sparse, symmetric, positive definite linear complementary systems, Appl. Math. Optim. 11 (1984), 77-95.
- [24] J. Mandel, Private Mitteilung, Oberwolfach 1984.

- [25] L.D. Marini, P. Pietra, Fixed-point algorithms for stationary flow in porous media, Submitted to Comput. Methods Appl. Mech. Engrg. .
- [26] P. Pietra, An up-wind method for a filtration problem, RAIRO (Num. An.) (4) 16 (1982), 463-481.
- [27] K. Stüben, U. Trottenberg, Multigrid methods: Fundamental algorithms, model problems, analysis and applications, 1-176 in [21].
- [28] R.S. Varga, Matrix iterative analysis, Englewood cliffs: Prentice-Hall 1962.

Lebenslauf

Am 16.6.1956 wurde ich als Sohn des kaufmännischen Angestellten Ewald Bollrath und seiner Frau Irmtraut Bollrath, geb. Stender, in Lünen geboren.

Nach dem Besuch der Marienschule in Lünen vom April 1963 bis zum November 1966 wechselte ich zum Gymnasium St. Christophorus in Werne a.d. Lippe. Dort erwarb ich am 19.6.1975 das Reifezeugnis.

Vom 1.7.1975 bis zum 30.9.1976 leistete ich meinen Grundwehrdienst bei der Bundeswehr.

Im Wintersemester 1976/77 begann ich das Studium der Mathematik mit Nebenfach Betriebswirtschaftslehre an der Ruhr-Universität Bochum und bestand dort am 27.9.1978 die Diplom-Vorprüfung. Am 29.10.1981 legte ich die Diplomprüfung ab.

Seit dem 2.11.1981 bin ich als wissenschaftlicher Mitarbeiter bei Herrn Prof. Dr. D. Braess an der Abteilung für Mathematik der Ruhr-Universität Bochum beschäftigt.

Bochum, im Februar 1985

